Postverlagsort Berlin

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT †
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR † A. E. INGHAM H. KNESER
W. MAGNUS O. PERRON G. PICKERT

72. BAND, 4. HEFT

(ABGESCHLOSSEN AM 22. JANUAR 1960)

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1960

Math. Z.

Die **MATHEMATISCHE ZEITSCHRIFT**

wurde im Jahre 1918 von *L. Lichtenstein* unter der Mitwirkung von *K. Knopp*, *E. Schmidt* und *I. Schur* gegründet und herausgegeben. Nach dem Tode Lichtensteins übernahm *K. Knopp* 1933 die Herausgabe. Die Schriftleitung ergänzte sich 1933 durch *E. Kamke* und *F. K. Schmidt*, 1936 durch *R. Nevanlinna* und 1950 durch *H. Wielandt*, der 1952 die Herausgabe übernahm.

Die Zeitschrift erscheint nach Maßgabe des eingehenden Materials. Der Preis des Bandes beträgt DM 96.—.

Die Mathematische Zeitschrift ist durch jede Buchhandlung zu beziehen.

Die Mathematische Zeitschrift dient der Pflege der reinen Mathematik, doch werden auch Beiträge aus den Gebieten der theoretischen Physik und Astronomie Aufnahme finden, soweit sie mathematisch von Interesse sind. Besprechungen, Aufgaben u. dgl. werden nicht zugelassen.

Grundsätzlich dürfen nur Arbeiten eingereicht werden, die vorher weder im Inland noch im Ausland veröffentlicht worden sind. Der Autor verpflichtet sich, sie auch nachträglich nicht an anderer Stelle zu publizieren. Mit der Annahme des Manuskriptes und seiner Veröffentlichung durch den Verlag geht das Verlagsrecht für alle Sprachen und Länder einschließlich des Rechts der fotomechanischen Wiedergabe oder einer sonstigen Vervielfältigung an den Verlag über. Jedoch wird gewerblichen Unternehmen für den innerbetrieblichen Gebrauch nach Maßgabe des zwischen dem Börsenverein des Deutschen Buchhandels e. V. und dem Bundesverband der Deutschen Industrie abgeschlossenen Rahmenabkommens die Anfertigung einer fotomechanischen Vervielfältigung gestattet. Wenn für diese Zeitschrift kein Pauschalabkommen mit dem Verlag vereinbart worden ist, ist eine Wertmarke im Betrage von DM 0.30 pro Seite zu verwenden. *Der Verlag läßt diese Beträge den Autorenverbänden zufließen.*

Die Mitarbeiter erhalten von ihren Arbeiten 75 Sonderdrucke unentgeltlich.

Manuskriptsendungen sind zu richten an die Mitglieder der unterzeichneten Redaktion. Die Herren Mitarbeiter werden im Interesse einer raschen Drucklegung gebeten, die Arbeiten in gut lesbarer Niederschrift einzureichen. Für den Text ist, wenn irgend möglich, Maschinenschrift, für die mathematischen Formeln jedoch nur Handschrift zu verwenden. Etwaige Abbildungen sind auf einem besonderen Blatt zu zeichnen. Die in den Formeln etwa vorkommenden griechischen oder deutschen Buchstaben (Fraktur) sowie andere der Gefahr von Verwechselungen ausgesetzte Zeichen, z.B. die Landauschen Symbole o und O, sind stets besonders (etwa mit verschiedenfarbigen Stiften) zu kennzeichnen. Die Fußnoten sind fortlaufend zu numerieren, bei Zitaten Erscheinungsjahr und Seitenzahlen anzugeben.

Wenn zu einer bereits gedruckten Arbeit eine Berichtigung nötig wird, wollen die Herren Verfasser dem Herausgeber hiervon unverzüglich Mitteilung machen. Die Berichtigungen sollen dann am Schlusse eines Bandes gesammelt aufgenommen werden.

REDAKTION DER MATHEMATISCHEN ZEITSCHRIFT:

E. Kamke, Tübingen, Frischlinstraße 27,

R. Nevanlinna, Helsinki (Finnland), Mäntytie 5 B

F. K. Schmidt, Heidelberg, Mathematisches Institut,

H. Wielandt, Tübingen, Ob dem Viehweidle 21 b.

Springer-Verlag OHG — Berlin-Wilmersdorf, Heidelberger Platz 3 — Heidelberg, Neuenheimer Landstraße 28/30. — Göttingen, Weender Straße 60.

---

72. Band **Inhalt** 4. Heft

Seite

Carlitz, L., Congruences (mod 2^r) for the Coefficients of the Jacobi Elliptic Functions 307

Specht, W., Freie Untergruppen der binären unimodularen Gruppe 319

Thompson, J. G., Normal p-complements for Finite Groups 332

Hodges, J. H., Some Determinantal Equations Over a Finite Field 355

Bagemihl, F., Planar sections of spatial sets 362

Kasch, F., und B. Volkmann, Metrische Sätze über transzendente Zahlen in P-adischen Körpern 367

Moser, J., The Order of a Singularity in Fuchs' Theory 379

Dem Andenken des Begründers dieser Zeitschrift

LEON LICHTENSTEIN

sind die folgenden Arbeiten gewidmet

CARLITZ, L.
Math. Zeitschr. 72, 307–318 (1960)

Congruences (mod 2^r) for the Coefficients of the Jacobi Elliptic Functions

By

L. CARLITZ

1. INTRODUCTION. Put

$$\operatorname{sn} x = \operatorname{sn}(x, k^2) = \sum_{m=0}^{\infty} A_{2m+1}(k^2)\,\frac{x^{2m+1}}{(2m+1)!}; \tag{1.1}$$

where $A_{2m+1}(k^2)$ is a polynomial in k^2 with integral coefficients. Also put

$$\frac{x}{\operatorname{sn} x} = \sum_{m=0}^{\infty} \beta_{2m}(k^2)\,\frac{x^{2m}}{(2m)!}, \tag{1.2}$$

where $\beta_{2m}(k^2)$ is a polynomial with rational coefficients. The writer has proved [2] that $A_n(k^2)$ satisfies

$$\sum_{s=0}^{r} (-1)^{r-s}\binom{r}{s} A_p^{(r-s)w/(p-1)} A_{n+sw} \equiv 0 \left(\operatorname{mod}(p^n, p^{er})\right), \tag{1.3}$$

where p is an *odd* prime,

$$p^{e-1}(p-1) \mid w \qquad (e \geqq 1) \tag{1.4}$$

and $r \geqq 1$. The result is to be interpreted in the following way: after expansion the coefficient of each power of k^2 in the left member $\equiv 0$.

In the next place the polynomial $\beta_{2m}(k^2)$ satisfies the congruence

$$\sum_{s=0}^{r} (-1)^{r-s}\binom{r}{s} A_p^{(r-s)w/(p-1)}\,\frac{\beta_{2m+sw}}{2m+sm} \equiv 0 \left(\operatorname{mod}(p^{2m-1}, p^{er})\right), \tag{1.5}$$

provided (1.4) holds and $p-1 \nmid 2m$; also we have

$$p\,\beta_{2m}(k^2) \equiv \begin{cases} -A_p^{2m/(p-1)} \pmod{p} & (p-1 \mid 2m) \\ 0 \pmod{p} & (\text{otherwise}) \end{cases} \tag{1.6}$$

and

$$\beta_{2m}(k^2) \equiv 0 \pmod{p^r} \qquad (p^r \mid 2m,\ p-1 \nmid 2m). \tag{1.7}$$

Results of a similar nature were obtained for the coefficients of the other Jacobi elliptic functions. Moreover the stated results were proved in a different way in [1].

It should be emphasized that in the above p is assumed odd throughout. In the present paper we examine the case $p=2$. The method of [1] does not seem to be applicable. Also the results are no longer quite so uniform. For example if we put

$$\operatorname{cn} x = \sum_{m=0}^{\infty} C_{2m}(k^2)\frac{x^{2m}}{(2m)!}, \tag{1.8}$$

then it was proved that

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}A_p^{(r-s)w/(p-1)}C_{2m+sw} \equiv 0 \pmod{(p^{2m}, p^{er})}, \tag{1.9}$$

where e, w satisfy (1.4). Now, on the other hand, we show that when $2^e \mid 2w$ and $r \geqq 1$, then

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}(1+k^2)^{(r-s)w}A_{n+2sw}(k^2) \equiv 0 \pmod{2^{er}}, \tag{1.10}$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}C_{n+2sw}(k^2) \equiv 0 \pmod{2^{er}}, \tag{1.11}$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}(1+k^2)^{(r-s)w}\beta_{n+2sw}(k^2) \equiv 0 \pmod{2^{er}}. \tag{1.12}$$

Note that powers of $1+k^2$ do not occur in the left member of (1.11). For the coefficients of $\operatorname{dn} x$, on the other hand, we get

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}k^{2(r-s)w}D_{n+2sw}(k^2) \equiv 0 \pmod{2^{er}}. \tag{1.13}$$

For other variants, see in particular (6.13), (6.14), (6.15), (7.12), (7.13) below.

2. We shall require a number of differentiation formulas. Let $D = d/dx$, so that

$$D \operatorname{sn} x = \operatorname{cn} x \operatorname{dn} x,$$

$$D^2 \operatorname{sn} x = -(1+k^2)\operatorname{sn} x - 2k^2 \operatorname{sn}^3 x,$$

so that

$$(D^2 + 1 + k^2)\operatorname{sn} x = -2k^2 \operatorname{sn}^3 x. \tag{2.1}$$

More generally

$$D \operatorname{sn}^n x = n \operatorname{sn}^{n-1} x \operatorname{cn} x \operatorname{dn} x,$$

$$D^2 \operatorname{sn}^n x = n(n-1)\operatorname{sn}^{n-2} x \operatorname{cn}^2 x \operatorname{dn}^2 x - n \operatorname{sn}^n x (1 + k^2 - 2k^2 \operatorname{sn}^2 x);$$

it follows that

$$(D^2 + 1 + k^2)\operatorname{sn}^n x$$
$$= n(n-1)\operatorname{sn}^{n-2} x \operatorname{cn}^2 x \operatorname{dn}^2 x - (n-1)(1+k^2)\operatorname{sn}^n x + 2nk^2 \operatorname{sn}^{n+2} x.$$

Thus

$$(2.2)\quad \begin{cases} (D^2+1+k^2)\,\mathrm{sn}^n x \\ \qquad = n(n-1)\,\mathrm{sn}^{n-2}x - (n^2-1)(1+k^2)\,\mathrm{sn}^n x + n(n+1)\,k^2\,\mathrm{sn}^{n+2}x; \end{cases}$$

if n is odd each numerical coefficient on the right side is even.

In the next place

$$D\,\mathrm{cn}\,x = -\,\mathrm{sn}\,x\,\mathrm{dn}\,x,$$

$$D^2\,\mathrm{cn}\,x = -(1-2k^2)\,\mathrm{cn}\,x - 2k^2\,\mathrm{cn}^3 x,$$

so that

$$(2.3)\quad (D^2+1)\,\mathrm{cn}\,x = 2k^2\,\mathrm{cn}\,x - 2k^2\,\mathrm{cn}^3 x.$$

More generally

$$D\,\mathrm{cn}^n x = -n\,\mathrm{cn}^{n-1}x\,\mathrm{sn}\,\mathrm{dn}\,x$$

$$D^2\,\mathrm{cn}^n x = n(n-1)\,\mathrm{cn}^{n-2}x\,\mathrm{sn}^2 x\,\mathrm{dn}^2 x - n\,\mathrm{cn}^n x(1-2k^2+2k^2\,\mathrm{cn}^2 x),$$

which gives

$$(2.4)\quad \begin{cases} (D^2+1)\,\mathrm{cn}^n x \\ \qquad = n(n-1)\,\mathrm{cn}^{n-2}x - (n^2-1)(1-2k^2)\,\mathrm{cn}^n x - n(n+1)\,k^2\,\mathrm{cn}^{n+2}x; \end{cases}$$

here also if n is odd, each coefficient on the right side is even.

In the same way for $\mathrm{dn}\,x$, we get

$$(2.5)\quad (D^2+k^2)\,\mathrm{dn}\,x = 2\,\mathrm{dn}\,x - 2\,\mathrm{dn}^3 x,$$

$$(2.6)\quad \begin{cases} (D^2+k^2)\,\mathrm{dn}^n x \\ \qquad = n(n-1)\,\mathrm{dn}^{n-2}x - (n^2-1)(k^2-2)\,\mathrm{dn}^n x - n(n+1)\,\mathrm{dn}^{n+2}x. \end{cases}$$

3. Now it is clear from repeated application of (2.1) and (2.2) that

$$(3.1)\quad (D^2+1+k^2)^r\,\mathrm{sn}\,x = 2^r \sum_{j=0}^{r} G_j(k^2)\,\mathrm{sn}^{2j+1}x,$$

where $G_j(k^2)$ is a polynomial in k^2 with integral coefficients. Thus, using (1.1), we get

$$(3.2)\quad \sum_{s=0}^{r} \binom{r}{s} (1+k^2)^{r-s} A_{2m+2s} \equiv 0 \pmod{2^r}.$$

Similarly, from (2.3) and (2.4),

$$(3.3)\quad (D^2+1)^r\,\mathrm{cn}\,x = 2^r \sum_{j=0}^{r} G_j(k^2)\,\mathrm{cn}^{2j+1}x,$$

where $G_j(k^2)$ is a polynomial in k^2 with integral coefficients (not necessarily the same as the one in (3.1)). Thus, using (1.8), we get

$$(3.4)\quad \sum_{s=0}^{r} \binom{r}{s} C_{2m+2s} \equiv 0 \pmod{2^r}.$$

Also, since (2.5) and (2.6) yield

$$(D^2+k^2)^r \operatorname{dn}^n x = 2^r \sum_{j=0}^{r} G_j(k^2) \operatorname{dn}^{2j+1} x, \tag{3.5}$$

it follows that

$$\sum_{s=0}^{r} \binom{r}{s} k^{2(r-s)} D_{2m+2s} \equiv 0 \pmod{2^r}, \tag{3.6}$$

where

$$\operatorname{dn} x = \sum_{m=0}^{\infty} D_{2m}(k^2) \frac{x^{2m}}{(2m)!}. \tag{3.7}$$

We remark that it follows from (2.1) that

$$A_{2n+3} + (1+k^2) A_{2n+1} \equiv 0 \pmod 4. \tag{3.7}$$

Also since $\operatorname{cn}^2 x \equiv \operatorname{dn}^2 x \equiv 1 \pmod 2$, it follows from (2.3) and (2.5) that

$$C_{2n+2} + C_{2n} \equiv 0 \pmod 4, \tag{3.8}$$

$$D_{2n+2} + k^2 D_{2n} \equiv 0 \pmod 4. \tag{3.9}$$

Thus we get

$$A_{2n+1} \equiv (-1)^n (1+k^2)^n \pmod 4, \tag{3.10}$$

$$C_{2n} \equiv (-1)^n \pmod 4, \tag{3.11}$$

$$D_{2n} \equiv (-1)^n k^{2n} \pmod 4. \tag{3.12}$$

4. To extend these results we require the following

LEMMA 1. *Let*

$$f(x) = f(x,u) = \sum_{n=0}^{\infty} f_n(u) \frac{x^n}{n!},$$

where the $f_n(u)$ are polynomials in the indeterminate u with rational coefficients that are integral (mod 2). *Also assume that*

$$(D^2 - \lambda)^r f(x) \equiv 0 \pmod{2^r} \tag{4.1}$$

for all $r \geqq 1$, where $\lambda = \lambda(u)$ is a polynomial with coefficients integral (mod 2). *Then we have*

$$(D^{2w} - \lambda^w)^r f(x) \equiv 0 \pmod{2^{er}}, \tag{4.2}$$

where $2^e \mid 2w$.

The statement

$$\sum_{n=0}^{\infty} g_n(u) \frac{x^n}{n!} \equiv 0 \pmod{2^r}$$

means

$$g_n(u) \equiv 0 \pmod{2^r} \qquad (n = 0, 1, 2, \ldots).$$

To prove the Lemma, we take

$$(4.3) \qquad D^{2w} - \lambda^w = \big((D^2 - \lambda) + \lambda\big)^w - \lambda^w = \sum_{j=1}^{w} \binom{w}{j} \lambda^{w-j} (D^2 - \lambda)^j,$$

so that

$$(4.4) \qquad (D^{2w} - \lambda^w) f(x) = \sum_{j=1}^{t} \binom{w}{j} \lambda^{w-j} (D^2 - \lambda)^j f(x).$$

Now by (4.1)

$$(D^2 - \lambda)^j f(x) \equiv 0 \pmod{2^j};$$

since $j!$ is divisible by 2^{j-1} at most, it follows that

$$\frac{1}{j!} (D^2 - \lambda)^j f(x) \equiv 0 \pmod 2 \qquad (j = 1, 2, 3, \ldots).$$

Since by hypothesis $2^{e-1} | w$, it is clear that (4.3) implies

$$(D^{2w} - \lambda^w) f(x) \equiv 0 \pmod{2^e}.$$

This is the case $r = 1$ of (4.2). For the general case we rewrite (4.3) in the form

$$D^{2w} - \lambda^w = w \sum_{j=1}^{w} \frac{1}{j} \binom{w-1}{j-1} \lambda^{w-j} (D^2 - \lambda)^j.$$

Raising both sides to the r-th power, we get

$$(D^{2w} - \lambda^w)^r f(x) = w^r \left\{ \sum_{j=1}^{w} \frac{1}{j} \binom{w-1}{j-1} \lambda^{w-j} (D^2 - \lambda)^j \right\}^r f(x).$$

Expanding the right member it is clear that each term $\equiv 0 \pmod{2^{er}}$. This completes the proof of the Lemma.

Another useful result is

LEMMA 2. *With the hypothesis of Lemma 1, let* $\lambda(u) \equiv \lambda_1(u) \pmod 2$, *where* $\lambda_1(u)$ *is another polynomial with coefficients integral* (mod 2). *Then we have*

$$(4.3) \qquad (D^{2w} - \lambda_1^w)^r f(x) \equiv 0 \pmod{2^{er}}.$$

Indeed if $\lambda(u) = \lambda_1(u) + 2\sigma(u)$, where σ has integral coefficients, then

$$(D^2 - \lambda_1)^r = (D^2 - \lambda + 2\sigma)^r = \sum_{s=0}^{r} \binom{r}{s} (2\sigma)^{r-s} (D^2 - \lambda)^s.$$

Thus if (4.1) holds it also holds when λ is replaced by λ_1. Now, applying Lemma 1, we obtain (4.3).

5. We now apply the Lemma of § 4 separately to (3.1), (3.3), (3.5) with $\lambda = -1 - k^2$, -1, $-k^2$, respectively. We may state

THEOREM 1. *Let* $r \geqq 1$ *and* $2^e | 2w$. *Then the coefficients* $A_n(k^2)$, $C_n(k^2)$, $D_n(k^2)$ *defined by* (1.1), (1.8), (3.7), *respectively, satisfy*

$$(5.1) \qquad \sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} (1 + k^2)^{(r-s)w} A_{n+2sw}(k^2) \equiv 0 \pmod{2^{er}},$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}C_{n+2sw}(k^2)\equiv 0 \pmod{2^{er}}, \tag{5.2}$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}k^{2(r-s)w}D_{n+2sw}(k^2)\equiv 0 \pmod{2^{er}}. \tag{5.3}$$

Indeed these results admit of an immediate generalization. Put

$$(\operatorname{sn} x)^j=\sum_{n=j}^{\infty}A_n^{(j)}(k^2)\frac{x^n}{n!},$$

$$(\operatorname{cn} x)^j=\sum_{n=0}^{\infty}C_n^{(j)}(k^2)\frac{x^n}{n!},$$

$$(\operatorname{dn} x)^j=\sum_{n=0}^{\infty}D_n^{(j)}(k^2)\frac{x^n}{n!};$$

then we have

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}(1+k^2)^{(r-s)w}A_{n+2sw}^{(j)}(k^2)\equiv 0 \pmod{2^{er}}, \tag{5.4}$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}C_{n+2sw}^{(j)}(k^2)\equiv 0 \pmod{2^{er}}, \tag{5.5}$$

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}D_{n+2sw}^{(j)}(k^2)\equiv 0 \pmod{2^{er}}, \tag{5.6}$$

where in (5.4) j is odd and positive, while in (5.5) and (5.6) it is enough to assume j odd.

For j even we return to the formulas (2.2), (2.4), (2.6). In view of the identities $\operatorname{cn}^2 x=1-\operatorname{sn}^2 x$, $\operatorname{dn}^2 x=1-k^2\operatorname{sn}^2 x$, it will suffice to consider (2.2), which we rewrite as

$$D^2\operatorname{sn}^n x=n(n-1)\operatorname{sn}^{n-2}x-n^2(1+k^2)\operatorname{sn}^n x+n(n+1)k^2\operatorname{sn}^{n+2}x.$$

This implies, for j even

$$A_{2r}^{(j)}\equiv 0 \pmod{2^r}. \tag{5.7}$$

Nothing is now gained by applying Lemma 1. The corresponding results for $C_{2r}^{(j)}$ and $D_{2r}^{(j)}$ are

$$C_{2r}^{(j)}\equiv 0 \pmod{2^r}, \tag{5.8}$$

$$D_{2r}^{(j)}\equiv 0 \pmod{2^r}. \tag{5.9}$$

We state

THEOREM 2. *For j odd, $A_n^{(j)}$, $C_n^{(j)}$, $D_n^{(j)}$ satisfy* (5.4), (5.5), (5.6), *where $2^e\mid 2w$; for j even, they satisfy* (5.7), (5.8), (5.9).

6. There are in all twelve Jacobi functions; it is convenient to separate them into two classes, namely sn, cn, dn, sc, sd, nc, nd, ed, dc in the first

class and ns, cs, ds in the second class — the notation is that of GLAISHER [4, p. 494]. The functions sn, cn, dn are covered by Theorem 1; while nc and nd are covered by the case $j = -1$ of (5.5) and (5.6). The remaining four functions in the first class can be handled by the method used above. It is however simpler to make use of the well-known formulas for the linear transformation of the Jacobi functions [3, p. 475], namely

$$\operatorname{sn}\left(k'x, \frac{k^2}{k^2-1}\right) = k' \operatorname{sd}(x, k^2), \tag{6.1}$$

$$\operatorname{cn}\left(k'x, \frac{k^2}{k^2-1}\right) = \operatorname{cd}(x, k^2), \tag{6.2}$$

$$\operatorname{dn}\left(k'x, \frac{k^2}{k^2-1}\right) = \operatorname{nd}(x, k^2), \tag{6.3}$$

$$\operatorname{sn}(ix, 1-k^2) = i \operatorname{sc}(x, k^2), \tag{6.4}$$

$$\operatorname{cn}(ix, 1-k^2) = \operatorname{nc}(x, k^2), \tag{6.5}$$

$$\operatorname{dn}(ix, 1-k^2) = \operatorname{dc}(x, k^2), \tag{6.6}$$

$$\operatorname{cn}\left(kx, \frac{1}{k^2}\right) = \operatorname{dn}(x, k^2). \tag{6.7}$$

We shall use the notation

$$\operatorname{sd}(x, k^2) = \sum_0^\infty \operatorname{sd}_n(k^2) \frac{x^n}{n!}, \tag{6.8}$$

and so on for the other functions not already introduced above. Now (6.1) implies

$$(1-k^2)^{\frac{1}{2}(n-1)} A_n\left(\frac{k^2}{k^2-1}\right) = \operatorname{sd}_n(k^2). \tag{6.9}$$

In (5.1) replace k^2 by $k^2/(k^2-1)$, so that

$$\sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} \left(\frac{1-2k^2}{1-k^2}\right)^{(r-s)w} A_{n+2sw}\left(\frac{k^2}{k^2-1}\right) \equiv 0 \pmod{2^{er}}.$$

Thus by Lemma 2 and (6.9) we get

$$\sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} \operatorname{sd}_{n+2sw} \equiv 0 \pmod{2^{er}}. \tag{6.10}$$

Similarly, since (6.2) yields

$$(1-k^2)^{\frac{1}{2}n} C_n\left(\frac{k^2}{k^2-1}\right) = \operatorname{cd}_n(k^2),$$

while (5.2) implies

$$\sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} C_{n+2sw}\left(\frac{k^2}{k^2-1}\right) \equiv 0 \pmod{2^{er}},$$

we find that

$$\sum_{s=0}^{r} (-1)^{r-s} \binom{r}{s} (1+k^2)^{(r-s)w} \operatorname{cd}_{n+2sw} \equiv 0 \pmod{2^{er}}. \tag{6.11}$$

In the same way we get from (6.3)

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}k^{2(r-s)w}\,\mathrm{nd}_{n+2sw}\equiv 0 \pmod{2^{er}}. \tag{6.12}$$

The formulas (6.4), (6.5), (6.6) give

$$\sum_{s=0}^{r}(-1)^{r-s+sw}\binom{r}{s}k^{2(r-s)w}\,\mathrm{sc}_{n+2sw}\equiv 0 \pmod{2^{er}}, \tag{6.13}$$

$$\sum_{s=0}^{r}(-1)^{r-s+sw}\binom{r}{s}\,\mathrm{nc}_{n+2sw}\equiv 0 \pmod{2^{er}}, \tag{6.14}$$

$$\sum_{s=0}^{r}(-1)^{r-s+sw}\binom{r}{s}(1+k^2)^{2(r-s)w}\,\mathrm{dc}_{n+2sw}\equiv 0 \pmod{2^{er}}. \tag{6.15}$$

Finally (6.7) implies the equivalence of (5.2) and (5.3).

We may state

THEOREM 3. *The coefficients of the functions* sd, cd, nd, sc, nc, dc, *satisfy* (6.10), (6.11), (6.12), (6.13), (6.14), (6.15), *respectively, where* $2^e \| 2w$ *and* $r \geqq 1$.

7. In addition to the notation defined in (1.2) we put

$$x\,\mathrm{cs}\,x=\sum_{m=0}^{\infty}\gamma_{2m}(k^2)\frac{x^{2m}}{(2m)!}, \tag{7.1}$$

$$x\,\mathrm{ds}\,x=\sum_{m=0}^{\infty}\delta_{2m}(k^2)\frac{x^{2m}}{(2m)!}. \tag{7.2}$$

To begin with, we note that the relation

$$\frac{d\,\mathrm{sn}\,x}{dx}=\mathrm{cn}\,x\,\mathrm{dn}\,x$$

implies

$$x=\sum_{n=0}^{\infty}\frac{\mathrm{sn}^{2n+1}x}{2x+1}\sum_{r+s=n}\frac{(\frac{1}{2})_r(\frac{1}{2})_s}{r!\,s!}k^{2r}=\sum_{n=0}^{\infty}\frac{\mathrm{sn}^{2n+1}x}{(2n+1)!}Q_n(k^2), \tag{7.3}$$

where

$$Q_n(k^2)=(2n)!\sum_{r+s=n}\frac{(\frac{1}{2})_r(\frac{1}{2})_s}{r!\,s!}k^{2r}.$$

The coefficients of $Q_n(k^2)$ are integral. Comparing (7.3) with (1.2) we conclude that the coefficients of $\beta_{2n}(k^2)$ are integral (mod 2). Since

$$\gamma_{2n}(k^2)=\sum_{r=0}^{n}\binom{2n}{2r}C_{n-r}(k^2)\beta_{2r}(k^2),$$

$$\delta_{2n}(k^2)=\sum_{r=0}^{n}\binom{2n}{2r}D_{n-r}(k^2)\beta_{2r}(k^2),$$

it follows that the coefficients of $\nu_{2n}(k^2)$ and $\delta_{2n}(k^2)$ are also integral (mod 2).

In the next place, we recall the duplication formula

$$\operatorname{sn} 2x = \frac{2 \operatorname{sn} x \operatorname{cn} x \operatorname{dn} x}{1 - k^2 \operatorname{sn}^4 x}.$$

From this we get

$$\frac{x}{\operatorname{sn} x} = \frac{2x}{\operatorname{sn} 2x} \frac{\operatorname{cn} x \operatorname{dn} x}{1 - k^2 \operatorname{sn}^4 x}. \tag{7.4}$$

If for brevity we put

$$\frac{\operatorname{cn} x \operatorname{dn}}{1 - k^2 \operatorname{sn}^4 x} = \sum_{m=0}^{\infty} \alpha_{2m}(k^2) \frac{x^{2m}}{(2m)!}, \tag{7.5}$$

then (7.4) and (1.2) imply

$$\beta_{2n}(k^2) = \sum_{r=0}^{n} \binom{2n}{2r} 2^{2r} \beta_{2r}(k^2) \alpha_{2n-2r}(k^2).$$

Since $\beta_{2r+1}(k^2) = \alpha_{2r+1}(k^2) = 0$, we may also write this in the form

$$\beta_n(k^2) = \sum_{r=0}^{n} \binom{n}{r} 2^r \beta_r(k^2) \alpha_{n-r}(k^2), \tag{7.6}$$

or, in symbolic form,

$$\beta^n = (2\beta + \alpha)^n. \tag{7.7}$$

Now consider

$$\begin{aligned}
(\beta^2 + 1 + k^2)^r &= \sum_{s=0}^{r} \binom{r}{s} (1 + k^2)^{r-s} \beta^{2s} \\
&= \sum_{s=0}^{r} \binom{r}{s} (1 + k^2)^{r-s} (2\beta + \alpha)^{2s} \\
&= \sum_{s=0}^{r} \binom{r}{s} (1 + k^2)^{r-s} \left(4(\beta^2 + \alpha\beta) + \alpha^2\right)^s \\
&= \sum_{s=0}^{r} \binom{r}{s} (1 + k^2)^{r-s} \sum_{j=0}^{s} \binom{s}{j} \alpha^{2(s-j)} 4^j (\beta^2 + \alpha\beta)^j \\
&= \sum_{j=0}^{r} \binom{r}{j} 4^j (\beta^2 + \alpha\beta)^j \sum_{s=j}^{r} \binom{r-j}{s-j} (1 + k^2)^{r-s} \alpha^{2(s-j)}.
\end{aligned}$$

Thus

$$(\beta^2 + 1 + k^2)^r = \sum_{j=0}^{r} \binom{r}{j} 4^j (\beta^2 + \alpha\beta)^j (\alpha^2 + 1 + k^2)^{r-j}. \tag{7.8}$$

Now the left member of (7.5) is equal to

$$\sum_{n=0}^{\infty} k^{2n} \operatorname{cn} x \operatorname{dn} x \operatorname{sn}^{4n} x = \frac{d}{dx} \sum_{n=0}^{\infty} k^{2n} \frac{\operatorname{sn}^{4n+1} x}{4n+1}.$$

Hence, using (5.4), we get

$$(\alpha^2 + 1 + k^2)^r \equiv 0 \pmod{2^r},$$

so that (7.8) yields

$$(\beta^2+1+k^2)^r \equiv 0 \pmod{2^r}. \tag{7.9}$$

In the next place

$$\beta^{2n}(\beta^2+1+k^2)^r = \big(\beta^2+1+k^2-(1+k^2)\big)^n(\beta^2+1+k^2)^r$$
$$= \sum_{s=0}^{n}(-1)^{n-s}(1+k^2)^{n-s}(\beta^2+1+k^2)^{r+s}.$$

In view of (7.9), this implies

$$\beta^{2n}(\beta^2+1+k^2)^r \equiv 0 \pmod{2^r}. \tag{7.10}$$

Now applying Lemma 1 to (7.10), we get

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}(1+k^2)^{(r-s)w}\beta_{2n+2sw}(k^2) \equiv 0 \pmod{2^{er}}, \tag{7.11}$$

where $2^e \mid 2w$.

As for the coefficients of cs x and ds x, we apply (6.4) and (6.1). We thus obtain

$$\sum_{s=0}^{r}(-1)^{r-s+sw}\binom{r}{s}k^{2(r-s)w}\gamma_{2n+2sw}(k^2) \equiv 0 \pmod{2^{er}}, \tag{7.12}$$

$$\sum_{s=0}^{r}(-1)^{r-s+sw}\binom{r}{s}\delta_{2n+2sw}(k^2) \equiv 0 \pmod{2^{er}}. \tag{7.13}$$

This completes the proof of

THEOREM 4. *The coefficients* $\beta_{2n}(k^2)$, $\gamma_{2n}(k^2)$, $\delta_{2n}(k^2)$ *defined by* (1.2), (7.1), (7.2), *respectively, satisfy* (7.11), (7.12), (7.13), *where* $2^e \mid 2w$ *and* $r \geqq 1$.

We remark that higher powers of ns, cs and ds can be handled without difficulty. If we put

$$\left(\frac{x}{\operatorname{sn} x}\right)^j = \sum_0^\infty \beta_{2n}^{(j)}(k^2)\frac{x^{2n}}{(2n)!},$$

$$(x\operatorname{cs} x)^j = \sum_0^\infty \gamma_{2n}^{(j)}(k^2)\frac{x^{2n}}{(2n)!},$$

$$(x\operatorname{ds} x)^j = \sum_0^\infty \delta_{2n}^{(j)}(k^2)\frac{x^{2n}}{(2n)!},$$

then the coefficients in $\beta_{2n}^{(j)}(k^2)$, $\gamma_{2n}^{(j)}(k^2)$, $\delta_{2n}^{(j)}(k^2)$ are integral (mod 2). We find that

$$\sum_{s=0}^{r}(-1)^{r-s}\binom{r}{s}(1+k^2)^{(r-s)w}\beta_{2n+2sw}^{(j)}(k^2) \equiv 0 \pmod{2^{er}} \tag{7.14}$$

for $2^e \mid 2w$ and j odd $\geqq 1$, while

$$\beta_{2r}^{(j)}(k^2) \equiv 0 \pmod{2^r} \tag{7.15}$$

for j even and $\geqq 0$. The results for $\gamma_{2n}^{(j)}(k^2)$ and $\delta_{2n}^{(j)}(k^2)$ are of a similar nature.

8. Since

$$\operatorname{sn}^4 x \equiv 0 \pmod 8,$$

(7.5) reduces to

$$\operatorname{cn} x \operatorname{dn} x \equiv \sum_{m=0}^{\infty} \alpha_{2m}(k^2) \frac{x^{2m}}{(2m)!} \pmod 8.$$

Comparing with (1.1) we get

$$\alpha_{2m}(k^2) \equiv A_{2m+1}(k^2) \pmod 8. \tag{8.1}$$

In particular it follows from (3.10) that

$$\alpha_{2m}(k^2) \equiv (-1)^m (1+k^2)^m \pmod 4. \tag{8.2}$$

Now from (7.7)

$$\beta^{2n} = \alpha^{2n} + \binom{2n}{2} (2\beta)^2 \alpha^{2n-2} + \cdots.$$

Thus if $2^e \mid 2n$ we get

$$\beta_{2n}(k^2) \equiv \alpha_{2n}(k^2) \pmod{2^{e+1}}; \tag{8.3}$$

indeed, using (8.2),

$$\beta_{2n}(k^2) \equiv \alpha_{2n}(k^2) + 4n(1+k^2)^n \pmod{2^{e+2}}. \tag{8.4}$$

In particular it is evident from (8.2) and (8.3) that

$$\beta_{2n}(k^2) \equiv (-1)^n (1+k^2)^n \pmod 4. \tag{8.5}$$

The corresponding formulas for $\gamma_{2n}(k^2)$ and $\delta_{2n}(k^2)$ may be obtained by applying (6.4) and (6.1). We get

$$(-1)^n \gamma_{2n}(k^2) \equiv \alpha_{2n}(1-k^2) + 4n k^{2n} \pmod{2^{e+2}}, \tag{8.6}$$

$$\delta_{2n}(k^2) \equiv (1+k^2)^n \alpha_{2n}\left(\frac{k^2}{k^2-1}\right) + 4n \pmod{2^{e+2}}, \tag{8.7}$$

$$\gamma_{2n}(k^2) \equiv (2-k^2)^n \pmod 4, \tag{8.8}$$

$$\delta_{2n}(k^2) \equiv (1-2k^2)^n \pmod 4. \tag{8.9}$$

As for higher powers, if we put

$$\left(\frac{\operatorname{cn} x \operatorname{dn} x}{1-k^2 \operatorname{sn}^4 x}\right)^j = \sum_{m=0}^{\infty} \alpha_{2m}^{(j)}(k^2) \frac{x^{2m}}{(2m)!} \qquad (j \geqq 1), \tag{8.10}$$

then (7.7) becomes

$$(\beta^{(j)})^n = (2\beta^{(j)} + \alpha^{(j)}). \tag{8.11}$$

It follows from (8.10) and (8.2) that

$$\alpha_{2n}^{(j)}(k^2) \equiv \begin{cases} (-1)^{\frac{1}{2}n(j+1)} (1+k^2)^n \pmod 4 & (j \text{ odd}) \\ 0 \pmod 4 & (j \equiv 0 \pmod 4) \\ 0 \pmod 4 & (j \equiv 2 \pmod 4, n > 1). \end{cases} \tag{8.12}$$

From (8.11) we obtain

$$\beta_{2n}^{(j)}(k^2) \equiv \alpha_{2n}^{(j)}(k^2) \pmod{2^{e+1}}, \tag{8.13}$$

where $2^e \mid 2n$; also using (8.12) a congruence (mod 2^{e+2}) is obtained. We note in particular that (8.12) and (8.13) yield

$$\beta_{2n}^{(j)}(k^2) \equiv \begin{cases} (-1)^{\frac{1}{2}n(j+1)}(1+k^2)^n \bmod 4 & (j \text{ odd}) \\ 0 \pmod 4 & \big(j \equiv 0 \pmod 4\big) \\ 0 \pmod 4 & \big(j \equiv 2 \pmod 4,\ n > 1\big). \end{cases} \tag{8.14}$$

Like results for $\gamma_{2n}^{(j)}(k^2)$ and $\delta_{2n}^{(j)}(k^2)$ are readily obtained.

References

[1] CARLITZ, L.: Arithmetic properties of elliptic functions. Math. Z. **64**, 425—434 (1956). — [2] CARLITZ, L.: Congruences connected with the power series expansions of the Jacobi elliptic functions. Duke Math. J. **20**, 1—12 (1953). — [3] FRICKE, R.: Die elliptischen Funktionen und ihre Anwendungen, Bd. 1. Leipzig u. Berlin 1916. — [4] WHITTAKER, E. T., and G. N. WATSON: Modern Analysis, 4. edit. Cambridge 1927.

Dept. of Math., Duke University, Durham, N.C. (U.S.A.)

(Eingegangen am 9. Februar 1959)

Specht, W.
Math. Zeitschr. 72, 319—331 (1960)

Freie Untergruppen der binären unimodularen Gruppe

Von
Wilhelm Specht

Die *binäre unimodulare Gruppe* $\mathfrak{U}_2$, d.h. die Gruppe aller zweireihigen komplexen Matrizen

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

mit der Determinante

$$\det M = \alpha\delta - \beta\gamma = 1$$

enthält, wie man weiß, freie Gruppen vom Range 2 und deshalb auch freie Gruppen von beliebigem endlichem und von abzählbar unendlichem Range als Untergruppen. Ein bekanntes[1]) Beispiel hierfür ist die von den beiden Matrizen

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

erzeugte freie Gruppe $\mathfrak{F}_2 = \{A, B\}$ vom Range 2, aus der nach bekanntem Verfahren durch die Matrizen

$$C_k = B^{-k} A B^k \qquad (-\infty < k < +\infty)$$

mit beliebigen ganzen Exponenten k eine freie Gruppe $\mathfrak{F}_\omega = \{\bigcup_k C_k\}$ abzählbar unendlichen Ranges gewonnen werden kann. Allerdings lassen sich bei diesem Beispiel die erzeugenden Matrizen C_k der Gruppe $\mathfrak{F}_\omega$ nicht leicht in einfacher geschlossener Gestalt angeben.

Ein in dieser Hinsicht bequemeres Beispiel bietet der

Satz 1. *Für jede (reelle oder komplexe) transzendente Zahl τ erzeugen die Matrizen*

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \textit{und} \quad V(\tau) = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \tag{1.1}$$

eine freie Gruppe

$$\mathfrak{H}_2(\tau) = \{U, V(\tau)\} \tag{1.2}$$

vom Range 2, die Matrizen

$$W_k(\tau) = U^{-k} V(\tau)\, U^k = \begin{pmatrix} 1 + k\tau & \tau \\ -k^2\tau & 1 - k\tau \end{pmatrix} \qquad (-\infty < k < +\infty) \tag{1.3}$$

[1]) Vgl. Magnus, W.: Untersuchungen über einige unendliche diskontinuierliche Gruppen. Math. Ann. **105**, 52—74 (1931).

mit beliebigen ganzen Exponenten k eine freie Gruppe

$$\mathfrak{H}_\omega(\tau) = \{\bigcup_k W_k(\tau)\} \tag{1.4}$$

abzählbar unendlichen Ranges.

Dieser Sachverhalt, dessen Beweis wir alsbald erbringen werden, legt die Parallelfrage nach der Struktur der durch die beiden Matrizen

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad V(\zeta) = \begin{pmatrix} 1 & \zeta \\ 0 & 1 \end{pmatrix}$$

erzeugten Gruppe $\mathfrak{H}_2(\zeta) = \{U, V(\zeta)\}$ nahe, in der ζ eine beliebige, von Null verschiedene, algebraische Zahl bezeichnet. Da eine vollständige Behandlung dieser Frage recht schwierig zu sein scheint, will ich mich in der vorliegenden Arbeit darauf beschränken, die Bedingungen anzugeben, unter denen die Gruppe $\mathfrak{H}_2(\zeta)$ eine freie Gruppe vom Range 2 ist.

Im Hinblick auf den Inhalt des Satzes 1 wollen wir der einfacheren Ausdrucksweise wegen eine von Null verschiedene komplexe Zahl ζ kurz als *Ausnahmewert* bezeichnen, wenn die durch sie bestimmte Gruppe $\mathfrak{H}_2(\zeta)$ keine freie Gruppe vom Range 2 ist, wenn also zwischen ihren Erzeugenden U und $V = V(\zeta)$ eine nichttriviale Relation besteht. Ohne Einschränkung darf dann angenommen werden, es bestehe eine *normierte Relation*

$$F_n(U, V) \equiv U^{a_0} V^{b_0} U^{a_1} V^{b_1} \dots U^{a_n} V^{b_n} = E \tag{2}$$

positiver *Länge* n mit ganzzahligen, von Null verschiedenen Exponenten $a_0, b_0, a_1, b_1, \dots, a_n, b_n$ und der zweireihigen Einheitsmatrix E. Überdies bezeichnen wir die komplexe Zahl ζ als *Ausnahmewert n-ter Stufe*, falls zwischen den Matrizen $U, V = V(\zeta)$ zwar eine normierte Relation der Länge n, aber keine normierte Relation geringerer Länge besteht.

Für einen beliebigen natürlichen Index $1 \leqq n < \infty$ bezeichnen wir ferner als *(abgeschlossenen) Kreisstreifen* $\overline{\mathfrak{C}}_n$ den kleinsten konvexen Bereich der komplexen Zahlenebene, der die beiden Kreisbereiche

$$|z+2| \leqq 2\cos\frac{\pi}{n+1} \quad \text{und} \quad |z-2| \leqq 2\cos\frac{\pi}{n+1}$$

enthält, also den (abgeschlossenen) Bereich aller komplexen Zahlen, die von der Strecke $\overline{\mathfrak{C}}_1 = \{-2 \leqq x \leqq 2\}$ auf der reellen Zahlengeraden höchstens den Abstand $2\cos\frac{\pi}{n+1}$ haben. Dann ist die Vereinigungsmenge $\mathfrak{C}_\omega = \bigcup_n \overline{\mathfrak{C}}_n$ ein *offener Kreisstreifen*, nämlich das (offene) Gebiet aller komplexen Zahlen, die von der Strecke $\mathfrak{C}_1$ einen kleineren Abstand als 2 besitzen (vgl. Figur).

Unter diesen Bezeichnungen gilt der merkwürdige

SATZ 2. *Jeder Ausnahmewert ζ n-ter Stufe ist eine algebraische Zahl höchstens vom Grade n, die dem abgeschlossenen Kreisstreifen $\overline{\mathfrak{C}}_n$ angehört.*

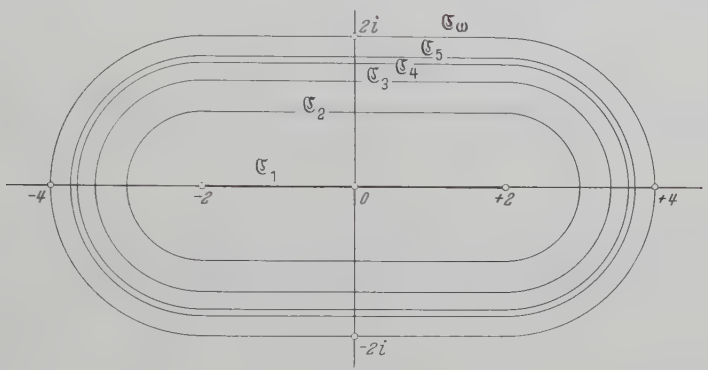
Unmittelbare Folgerung dieses Ergebnisses ist der

Satz 2*. *Für jede außerhalb des offenen Kreisstreifens* $\mathfrak{C}_\omega$ *liegende algebraische Zahl* ζ *erzeugen die Matrizen*

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \textit{und} \quad V(\zeta) = \begin{pmatrix} 1 & \zeta \\ 0 & 1 \end{pmatrix} \tag{3.1}$$

eine freie Gruppe $\mathfrak{H}_2(\zeta)$ *vom Range 2, die Matrizen*

$$W_k(\zeta) = U^{-k} V(\zeta)\, U^k = \begin{pmatrix} 1 + k\zeta & \zeta \\ -k^2\zeta & 1 - k\zeta \end{pmatrix} \qquad (-\infty < k < \infty) \tag{3.2}$$



mit beliebigen ganzen Exponenten k *eine freie Gruppe*

$$\mathfrak{H}_\omega(\zeta) = \{\bigcup_k W_k(\zeta)\} \tag{3.3}$$

abzählbar unendlichen Ranges.

Als Kennzeichnung der Ausnahmewerte kann die nachstehende Aussage dienen:

Satz 3. *Zu einer beliebig vorgegebenen Serie*

$$b_0, a_1, b_1, a_2, b_2, \ldots, a_n, b_n \tag{4.1}$$

von Null verschiedener ganzer Zahlen beliebiger Länge $1 \leqq n < \infty$ *bilde man für* $1 \leqq \nu < n$ *rekursiv die Polynomserie*

$$\psi_{\nu+1}(z) = \left(z + \frac{1}{b_\nu a_{\nu+1}} + \frac{1}{a_{\nu+1} b_{\nu+1}}\right)\psi_\nu(z) - \frac{1}{a_\nu b_\nu^2 a_{\nu+1}}\psi_{\nu-1}(z) \tag{4.2}$$

mit den Anfangsgliedern

$$\psi_0(z) = 1; \quad \psi_1(z) = z + \frac{1}{b_0 a_1} + \frac{1}{a_1 b_1}. \tag{4.3}$$

Dann ist jede von Null verschiedene Nullstelle ζ *des letzten Polynoms* $\psi_n(z)$ *(vom Grade* n*) ein Ausnahmewert (aus dem abgeschlossenen Kreisstreifen* $\overline{\mathfrak{C}}_n$*).*

Zwischen den Erzeugenden

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \textit{und} \quad V = V(\zeta) = \begin{pmatrix} 1 & \zeta \\ 0 & 1 \end{pmatrix}$$

der durch ζ bestimmten Gruppe $\mathfrak{H}_2(\zeta)$ besteht unter der Abkürzung

$$S_n \equiv S_n(U, V) \equiv V^{b_0} U^{a_1} V^{b_1} U^{a_2} V^{b_2} \dots U^{a_n} V^{b_n} \tag{5.1}$$

die normierte Relation

$$F_{4n}(U, V) \equiv U^{-1} S_n U S_n^{-1} U S_n U^{-1} S_n^{-1} = E \tag{5.2}$$

der Länge $4n$.

Umgekehrt existiert zu jedem Ausnahmewert ζ n-ter Stufe eine Serie

$$b_0, a_1, b_1, a_2, b_2, \dots, a_n, b_n$$

von Null verschiedener ganzer Zahlen, derart daß ζ Nullstelle des durch diese Serie bestimmten Polynoms $\psi_n(z)$ (vom Grade n) ist.

1.

Mit einer komplexen Veränderlichen z bilden wir für ein beliebiges Paar a, b von Null verschiedener ganzer Zahlen die Matrizen

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad V = V(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix},$$

deren Potenzen

$$U^a = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} \quad \text{bzw.} \quad V^b = \begin{pmatrix} 1 & bz \\ 0 & 1 \end{pmatrix}$$

und aus diesen das Produkt

$$R_0\left(\begin{matrix} a \\ b \end{matrix} \middle| z\right) = U^a V^b = \begin{pmatrix} 1 & bz \\ a & 1 + abz \end{pmatrix}. \tag{6.1}$$

Für eine beliebige Folge

$$a_0, b_0, a_1, b_1, a_2, b_2, \dots, a_n, b_n, a_{n+1}, b_{n+1}, \dots$$

von Null verschiedener ganzer Zahlen erklären wir dann die Matrixfolge

$$\begin{cases} R_n(z) = R_n\left(\begin{matrix} a_0 & a_1 & \dots & a_n \\ b_0 & b_1 & \dots & b_n \end{matrix} \middle| z\right) \\ \quad = R_0\left(\begin{matrix} a_0 \\ b_0 \end{matrix} \middle| z\right) R_0\left(\begin{matrix} a_1 \\ b_1 \end{matrix} \middle| z\right) \dots R_0\left(\begin{matrix} a_n \\ b_n \end{matrix} \middle| z\right) \\ \quad = U^{a_0} V^{b_0} U^{a_1} V^{b_1} \dots U^{a_n} V^{b_n} \qquad (\text{für } 0 \leqq n < \infty). \end{cases} \tag{6.2}$$

Offensichtlich sind die Koeffizienten der Matrix $R_n(z)$ ganzzahlige Polynome in der Veränderlichen z, weshalb die Matrix $R_n(z)$ in der Gestalt

$$R_n(z) = \begin{pmatrix} \alpha_n(z) & z\beta_n(z) \\ \gamma_n(z) & \delta_n(z) \end{pmatrix} \tag{6.3}$$

angenommen werden kann. In der Tat gilt

$$R_0(z) = \begin{pmatrix} 1 & b_0 z \\ a_0 & 1 + a_0 b_0 z \end{pmatrix} = \begin{pmatrix} \alpha_0(z) & z\beta_0(z) \\ \gamma_0(z) & \delta_0(z) \end{pmatrix},$$

also

$$\alpha_0(z) = 1; \qquad \beta_0(z) = b_0$$
$$\gamma_0(z) = a_0; \qquad \delta_0(z) = 1 + a_0 b_0 z,$$

ferner

$$\begin{aligned} R_{n+1}(z) &= R_{n+1}\left(\begin{matrix} a_0 & a_1 & \cdots & a_n & a_{n+1} \\ b_0 & b_1 & \cdots & b_n & b_{n+1} \end{matrix} \middle| \, z\right) \\ &= R_n\left(\begin{matrix} a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_n \end{matrix} z\right) R_0\left(\begin{matrix} a_{n+1} \\ b_{n+1} \end{matrix} \middle| \, z\right) \\ &= \begin{pmatrix} \alpha_n(z) & z\beta_n(z) \\ \gamma_n(z) & \delta_n(z) \end{pmatrix} \begin{pmatrix} 1 & z b_{n+1} \\ a_{n+1} & 1 + a_{n+1} b_{n+1} z \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{n+1}(z) & z\beta_{n+1}(z) \\ \gamma_{n+1}(z) & \delta_{n+1}(z) \end{pmatrix}, \end{aligned}$$

also

$$\begin{aligned} \alpha_{n+1}(z) &= \alpha_n(z) + a_{n+1} z \beta_n(z), \\ \beta_{n+1}(z) &= b_{n+1}\alpha_n(z) + (1 + a_{n+1} b_{n+1} z)\beta_n(z) \\ &= \beta_n(z) + b_{n+1}\left(\alpha_n(z) + a_{n+1} z \beta_n(z)\right) \\ &= \beta_n(z) + b_{n+1}\alpha_{n+1}(z), \\ \gamma_{n+1}(z) &= \gamma_n(z) + a_{n+1}\delta_n(z), \\ \delta_{n+1}(z) &= b_{n+1} z \gamma_n(z) + (1 + a_{n+1} b_{n+1} z)\delta_n(z) \\ &= \delta_n(z) + b_{n+1} z\left(\gamma_n(z) + a_{n+1}\delta_n(z)\right) \\ &= \delta_n(z) + b_{n+1} z \gamma_{n+1}(z). \end{aligned}$$

Für die in den Matrizen $R_n(z)$ auftretenden ganzzahligen Polynome gelten demnach die gekoppelten Rekursionsgleichungen

(7.1) $\alpha_{n+1}(z) = \alpha_n(z) + a_{n+1} z \beta_n(z); \qquad \beta_{n+1}(z) = \beta_n(z) + b_{n+1}\alpha_{n+1}(z);$

(7.2) $\gamma_{n+1}(z) = \gamma_n(z) + a_{n+1}\delta_n(z); \qquad \delta_{n+1}(z) = \delta_n(z) + b_{n+1} z \gamma_{n+1}(z)$

mit den Anfangsgliedern

(8.1) $\alpha_0(z) = 1; \qquad \beta_0(z) = b_0,$

(8.2) $\gamma_0(z) = a_0; \qquad \delta_0(z) = 1 + a_0 b_0 z.$

Überdies besteht für jede Matrix $R_n(z)$ die Identität

(9) $$\det R_n(z) \equiv \alpha_n(z)\,\delta_n(z) - z\beta_n(z)\,\gamma_n(z) = 1.$$

Die Polynome $\alpha_0(z), \beta_0(z), \gamma_0(z)$ sind vom Grade 0, das Polynom $\delta_0(z)$ ist vom Grade 1. Unter der Induktionsannahme, die Polynome $\alpha_n(z), \beta_n(z), \gamma_n(z)$ seien vom Grade n, das Polynom $\delta_n(z)$ sei vom Grade $n+1$, entnimmt man

den Rekursionsgleichungen unmittelbar die Folgerung, daß jedes der Polynome $\alpha_{n+1}(z), \beta_{n+1}(z), \gamma_{n+1}(z)$ vom Grade $n+1$, das Polynom $\delta_{n+1}(z)$ vom Grade $n+2$ ist. Mithin ist in der Matrix $R_n(z)$ jedes der Polynome $\alpha_n(z)$, $\beta_n(z), \gamma_n(z)$ vom Grade n, das Polynom $\delta_n(z)$ vom Grade $n+1$.

Infolgedessen ist auch für jede (reelle oder komplexe) transzendente Zahl τ und jede Serie $a_0, b_0, a_1, b_1, \ldots, a_n, b_n$ von Null verschiedener ganzer Zahlen die Matrix

$$R_n(\tau) = R_n\left(\begin{matrix} a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_n \end{matrix} \middle| \tau\right) = \begin{pmatrix} \alpha_n(\tau) & \tau\beta_n(\tau) \\ \gamma_n(\tau) & \delta_n(\tau) \end{pmatrix}$$

von der Einheit E verschieden. Damit ist die erste Aussage des Satzes 1 bewiesen. Die zweite Aussage ergibt sich nun aus den Darstellungen

$$W_k(\tau) = U^{-k}V(\tau)\,U^k = \begin{pmatrix} 1+k\tau & \tau \\ -k^2\tau & 1-k\tau \end{pmatrix} \qquad (-\infty < k < +\infty)$$

in bekannter Weise nach allgemeinen Überlegungen[2]).

2.

Auf Grund des Satzes 1 ist ein Ausnahmewert n-ter Stufe eine algebraische Zahl ζ, für die zwischen den erzeugenden Matrizen

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad V = V(\zeta) = \begin{pmatrix} 1 & \zeta \\ 0 & 1 \end{pmatrix}$$

der Gruppe $\mathfrak{H}_2(\zeta)$ eine normierte Relation

$$F_n(U, V) \equiv U^{a_0}V^{b_0}U^{a_1}V^{b_1}\ldots U^{a_n}V^{b_n} = E \tag{10.1}$$

der Länge n besteht. Erklären wir demgemäß für die Serie

$$a_0, b_0, a_1, b_1, \ldots, a_n, b_n$$

von Null verschiedener ganzer Exponenten entsprechend den Rekursionsgleichungen (7) mit den Anfangsgliedern (8) die Matrix

$$R_n(z) = R_n\left(\begin{matrix} a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_n \end{matrix} \middle| z\right) = \begin{pmatrix} \alpha_n(z) & z\beta_n(z) \\ \gamma_n(z) & \delta_n(z) \end{pmatrix},$$

so gilt auf Grund der Voraussetzung die Gleichung

$$R_n(\zeta) = E \quad \text{oder} \quad \begin{pmatrix} \alpha_n(\zeta) & \zeta\beta_n(\zeta) \\ \gamma_n(\zeta) & \delta_n(\zeta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{10.2}$$

insbesondere also die Gleichung

$$\beta_n(\zeta) = 0. \tag{11}$$

Da $\beta_n(z)$ ein ganzzahliges Polynom vom Grade n ist, ist der Ausnahmewert ζ n-ter Stufe in der Tat eine algebraische Zahl höchstens vom Grade n. Zum

[2]) Vgl. etwa SPECHT, W.: Gruppentheorie, Abschnitt 1.2.4, Beispiel 3. Berlin: Springer 1956.

Beweise des Satzes 2 bleibt noch nachzuweisen, daß die Nullstellen des Polynoms $\beta_n(z)$ dem abgeschlossenen Kreisstreifen $\overline{\mathfrak{S}}_n$ der komplexen Zahlenebene angehören.

Aus den Rekursionsgleichungen (7) und (8) entnehmen wir zunächst die Darstellungen

$$\alpha_1(z) = \alpha_0(z) + a_1 z \beta_0(z) = 1 + b_0 a_1 z,$$
$$\beta_1(z) = \beta_0(z) + b_1 \alpha_1(z) = b_0 + b_1 + b_0 a_1 b_1 z.$$

Aus den Gln. (7.1) erhalten wir weiter

$$\alpha_{\nu+1}(z) = \frac{\beta_{\nu+1}(z) - \beta_\nu(z)}{b_{\nu-1}}; \quad \alpha_\nu(z) = \frac{\beta_\nu(z) - \beta_{\nu-1}(z)}{b_\nu} \quad \text{für } 1 \leqq \nu < n$$

und

$$\frac{\beta_{\nu+1}(z) - \beta_\nu(z)}{b_{\nu+1}} = \frac{\beta_\nu(z) - \beta_{\nu-1}(z)}{b_\nu} + a_{\nu+1} z \beta_\nu(z) \quad \text{für } 1 \leqq \nu < n.$$

Mithin genügt die Polynomserie

$$\beta_0(z),\ \beta_1(z),\ \beta_2(z),\ \ldots,\ \beta_n(z)$$

der zweigliedrigen Rekursionsgleichung

$$\frac{\beta_{\nu+1}(z)}{a_{\nu+1} b_{\nu+1}} = \left(z + \frac{1}{b_\nu a_{\nu+1}} + \frac{1}{a_{\nu+1} b_{\nu+1}}\right) \beta_\nu(z) - \frac{\beta_{\nu-1}(z)}{b_\nu a_{\nu+1}} \tag{12.1}$$

mit den Anfangsgliedern

$$\beta_0(z) = b_0; \quad \beta_1(z) = b_0 a_1 b_1 z + b_1 + b_0. \tag{12.2}$$

Jedes Polynom $\beta_\nu(z)$ dieser Serie besitzt die gleichen Nullstellen wie das zugehörige normierte Polynom

$$\psi_\nu(z) = \frac{\beta_\nu(z)}{b_0 a_1 b_1 a_2 b_2 \ldots a_\nu b_\nu} \quad \text{für } 1 \leqq \nu \leqq n. \tag{13.1}$$

Aus den Gln. (12) erhält man für die Polynomserie

$$\psi_0(z),\ \psi_1(z),\ \psi_2(z),\ \ldots,\ \psi_n(z)$$

die Rekursionsgleichung

$$\psi_{\nu+1}(z) = \left(z + \frac{1}{b_\nu a_{\nu+1}} + \frac{1}{a_{\nu+1} b_{\nu+1}}\right) \psi_\nu(z) - \frac{1}{a_\nu b_\nu^2 a_{\nu+1}} \psi_{\nu-1}(z) \text{ für } 1 \leqq \nu < n \tag{13.2}$$

mit den Anfangsgliedern

$$\psi_0(z) = 1; \quad \psi_1(z) = z + \frac{1}{b_0 a_1} + \frac{1}{a_1 b_1}, \tag{13.3}$$

in Übereinstimmung mit der letzten Aussage des Satzes 3, da die Polynome $\psi_\nu(z)$ unabhängig vom ersten Exponenten a_0 sind.

Setzen wir nun noch zur Abkürzung

$$c_\nu = \frac{1}{a_\nu b_\nu} \quad \text{und} \quad d_\nu = \frac{1}{b_{\nu-1} a_\nu} \quad \text{für } 1 \leqq \nu \leqq n, \tag{14.1}$$

sowie

$$(14.2)\qquad c_\nu^* = c_\nu \frac{\sin\frac{(\nu+1)\pi}{n+1}}{\sin\frac{\nu\pi}{n+1}} \quad \text{und} \quad d_\nu^* = d_\nu \frac{\sin\frac{(\nu-1)\pi}{n+1}}{\sin\frac{\nu\pi}{n+1}} \qquad \text{für } 1 \leqq \nu \leqq n,$$

so gilt

$$(14.3)\qquad c_\nu d_{\nu+1} = c_\nu^* d_{\nu+1}^* \qquad \text{für } 1 \leqq \nu < n;$$

überdies bestehen die Beschränkungen

$$(14.4)\qquad -2 \leqq c_\nu + d_\nu \leqq +2,$$

$$(14.5)\qquad 0 < |c_\nu^*| + |d_\nu^*| \leqq \frac{\sin\frac{(\nu+1)\pi}{n+1}}{\sin\frac{\nu\pi}{n+1}} + \frac{\sin\frac{(\nu-1)\pi}{n+1}}{\sin\frac{\nu\pi}{n+1}} = 2\cos\frac{\pi}{n+1}$$

für $1 \leqq \nu \leqq n$.

Da die Rekursionsgleichung der Polynomserie $\{\psi_\nu(z)\}$ hiernach auch in die Gestalt

$$(15.1)\qquad \psi_{\nu+1}(z) = (z + c_{\nu+1} + d_{\nu+1})\,\psi_\nu(z) - c_\nu^* d_{\nu+1}^* \psi_{\nu-1}(z) \qquad \text{für } 1 \leqq \nu < n,$$

$$(15.2)\qquad \psi_0(z) = 1; \quad \psi_1(z) = z + c_1 + d_1$$

gebracht werden kann, ist das letzte Polynom $\psi_n(z)$ dieser Serie das charakteristische Polynom[3])

$$(15.3)\qquad \psi_n(z) = \det(zE - M_n)$$

der n-reihigen Matrix

$$(15.4)\qquad M_n = \begin{pmatrix} -(c_1+d_1) & c_1^* & 0 & \dots & 0 & 0 \\ d_2^* & -(c_2+d_2) & c_2^* & \dots & 0 & 0 \\ 0 & d_3^* & -(c_3+d_3) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -(c_{n-1}+d_{n-1}) & c_{n-1}^* \\ 0 & 0 & 0 & \dots & d_n^* & -(c_n+d_n) \end{pmatrix}.$$

Die gewünschte Aussage über die Lage der Nullstellen des Polynoms $\beta_n(z)$ (oder des Polynoms $\psi_n(z)$) gewinnt man nun fast unmittelbar aus dem folgenden bekannten[4])

LEMMA. *Bildet man aus den Elementen einer n-reihigen komplexen Matrix*

$$A = (a_{\varkappa\lambda}) \qquad (1 \leqq \varkappa, \lambda \leqq n)$$

die Werte

$$s_\varkappa = -|a_{\varkappa\varkappa}| + \sum_\lambda |a_{\varkappa\lambda}| \qquad \textit{für } 1 \leqq \varkappa \leqq n,$$

[3]) SPECHT, W.: Die Lage der Nullstellen eines Polynoms. III. Math. Nachr. **16**, 369–389 (1957).

[4]) Vgl. BRAUER, A.: Limits for the characteristic roots of a matrix. Duke Math. J. **13**, 387–395 (1946).

so gehört jeder Eigenwert ω der Matrix A (wenigstens) einem der n Kreisbereiche

$$|z - a_{\varkappa\varkappa}| \leqq s_\varkappa \qquad \text{für } 1 \leqq \varkappa \leqq n$$

der komplexen Zahlenebene an.

Nach diesem Lemma gehört jeder Eigenwert der Matrix M_n, also jede Nullstelle ζ des Polynoms $\psi_n(z)$ (oder des Polynoms $\beta_n(z)$) wenigstens einem der Kreisbereiche

$$|z + c_\nu + d_\nu| \leqq |c_\nu^*| + |d_\nu^*| \qquad \text{für } 1 \leqq \nu \leqq n,$$

also auf Grund der Beschränkungen (14.4) und (14.5) auch wenigstens einem der Kreisbereiche

$$|z - t| \leqq 2\cos\frac{\pi}{n+1} \qquad \text{mit } -2 \leqq t \leqq +2$$

an. Die Vereinigung dieser Kreisbereiche ist aber, wie man sofort erkennt, der in der Einleitung erklärte abgeschlossene Kreisstreifen $\overline{\mathfrak{C}}_n$ der komplexen Zahlenebene.

Damit ist der Satz 2 vollständig bewiesen. Liegt eine algebraische Zahl ζ außerhalb des offenen Kreisstreifens $\mathfrak{C}_\omega$, also außerhalb jedes abgeschlossenen Kreisstreifens $\overline{\mathfrak{C}}_n$, so ist sie gewiß kein Ausnahmewert; folglich gilt auch der Satz 2*.

3.

Nun sei zur Vervollständigung des Beweises des Satzes 3 umgekehrt eine beliebige Serie

$$b_0, a_1, b_1, a_2, b_2, \ldots, a_n, b_n$$

von Null verschiedener ganzer Zahlen vorgegeben und den Gln. (4.1), (4.2) des Satzes 3 entsprechend die Serie

$$\psi_0(z),\ \psi_1(z),\ \psi_2(z),\ \ldots,\ \psi_n(z)$$

rationalzahliger Polynome durch die Rekursionsgleichung

$$\psi_{\nu+1}(z) = \left(z + \frac{1}{b_\nu a_{\nu-1}} + \frac{1}{a_{\nu-1} b_{\nu-1}}\right)\psi_\nu(z) - \frac{1}{a_\nu b_\nu^2 a_{\nu+1}}\psi_{\nu-1}(z) \qquad \text{für } 1 \leqq \nu < n$$

mit den Anfangsgliedern

$$\psi_0(z) = 1; \qquad \psi_1(z) = z + \frac{1}{b_0 a_1} + \frac{1}{a_1 b_1}$$

erklärt. Bildet man dann mit einer beliebig ergänzten Exponentenserie

$$a_0, b_0, a_1, b_1, a_2, b_2, \ldots, a_n, b_n$$

aus den Matrizen

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \text{und} \qquad V(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}$$

das Matrixprodukt

$$\begin{aligned} R_n(z) &= R_n\begin{pmatrix} a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_n \end{pmatrix} z \\ &= U^{a_0} V^{b_0}(z)\, U^{a_1} V^{b_1}(z) \ldots U^{a_n} V^{b_n}(z) \\ &= \begin{pmatrix} \alpha_n(z) & z\,\beta_n(z) \\ \gamma_n(z) & \delta_n(z) \end{pmatrix} \end{aligned}$$

mit der Determinante

$$\det R_n(z) \equiv \alpha_n(z)\,\delta_n(z) - z\,\beta_n(z)\,\gamma_n(z) \equiv 1\,,$$

so stehen die Polynome $\psi_n(z)$ und $\beta_n(z)$ in dem durch die Gl. (13.1) gegebenen Zusammenhang:

$$\psi_n(z) = \frac{\beta_n(z)}{b_0 a_1 b_1 a_2 b_2 \dots a_n b_n}\,.$$

Für eine beliebig ausgewählte, von Null verschiedene Nullstelle ζ des Polynoms $\psi_n(z)$ (aus dem abgeschlossenen Kreisstreifen $\overline{\mathfrak{C}}_n$) erhalten wir dann

$$R_n(\zeta) = \begin{pmatrix} \alpha_n(\zeta) & 0 \\ \gamma_n(\zeta) & \delta_n(\zeta) \end{pmatrix}$$

mit der Determinante

$$\det R_n(\zeta) = \alpha_n(\zeta)\,\delta_n(\zeta) = 1\,.$$

Daß hierbei die normierte Relation

$$R_n(\zeta) = E$$

erfüllt ist, kann natürlich allgemein nicht erwartet werden. Setzen wir indes

$$\begin{aligned} S_n = S_n(\zeta) &= V^{b_0} U^{a_1} V^{b_1} \dots U^{a_n} V^{b_n} = U^{-a_0} R_n(\zeta) \\ &= \begin{pmatrix} 1 & 0 \\ -a_0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_n(\zeta) & 0 \\ \gamma_n(\zeta) & \delta_n(\zeta) \end{pmatrix} = \begin{pmatrix} \alpha_n(\zeta) & 0 \\ \sigma_n(\zeta) & \delta_n(\zeta) \end{pmatrix} \end{aligned}$$

mit

$$\sigma_n(\zeta) = \gamma_n(\zeta) - a_0 \alpha_n(\zeta)\,,$$

so findet man durch eine einfache Rechnung

$$\begin{aligned} U^{-1} S_n U S_n^{-1} &= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \alpha_n(\zeta) & 0 \\ \sigma_n(\zeta) & \delta_n(\zeta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \delta_n(\zeta) & 0 \\ -\sigma_n(\zeta) & \alpha_n(\zeta) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \delta_n^2(\zeta) - 1 & 1 \end{pmatrix}, \end{aligned}$$

also

$$U^{-1} S_n U S_n^{-1} \cdot U = U \cdot U^{-1} S_n U S_n^{-1}$$

und damit die normierte Relation

$$F_{4n}(U, V) = U^{-1} S_n U S_n^{-1} U S_n U^{-1} S_n^{-1} = E\,.$$

Mithin ist die Nullstelle ζ des Polynoms $\psi_n(z)$ ein Ausnahmewert höchstens $4n$-ter Stufe. Damit ist auch der Satz 3 vollständig bewiesen.

4.

Nach Satz 1 erzeugen die beiden Matrizen

$$U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad V = V(4) = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$$

eine freie Gruppe $\mathfrak{H}_2(4)$ vom Range 2; das gleiche gilt infolgedessen für die beiden Matrizen

$$X = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}; \qquad Y = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Mithin besteht auch der

Satz 4. *Die beiden unimodularen Matrizen*

$$X = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad \textit{und} \quad Y = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

erzeugen eine freie Gruppe

$$\mathfrak{G}_2 = \{X, Y\}$$

vom Range 2, die Matrizen

$$Z_k = \begin{pmatrix} 1 + 4k & -4k^2 \\ 4 & 1 - 4k \end{pmatrix}$$

mit beliebigem ganzzahligem Index k eine freie Gruppe

$$\mathfrak{G}_\omega = \{\bigcup_k Z_k\}$$

von abzählbar unendlichem Rang.

Die angegebenen Matrizen Z_k sind nämlich die Transformierten

$$Z_k = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2k \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2k \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+4k & -4k^2 \\ 4 & 1-4k \end{pmatrix};$$

daher folgt die zweite Behauptung des Satzes aus einem allgemeingültigen Sachverhalt.

Erklären wir zu den beiden besonderen Serien

$$b_0 = b_1 = b_2 = \cdots = b_n = 1; \qquad a_1 = a_2 = \cdots = a_n = \varepsilon = \pm 1$$

ganzer Zahlen die Polynomserien

$$\psi_0^\varepsilon(z),\ \psi_1^\varepsilon(z),\ \ldots,\ \psi_n^\varepsilon(z)$$

entsprechend den Gln. (13.1) und (13.2) durch die Rekursionsgleichungen

$$\psi_{\nu+1}^\varepsilon(z) = (z + 2\varepsilon)\,\psi_\nu^\varepsilon(z) - \psi_{\nu-1}^\varepsilon(z) \qquad \text{für } 1 \leqq \nu < n$$

mit den Anfangsgliedern

$$\psi_0^\varepsilon(z) = 1; \qquad \psi_1^\varepsilon(z) = z + 2\varepsilon,$$

so lassen sich die Nullstellen des letzten Polynoms $\psi_n^\varepsilon(z)$ leicht angeben. Setzen wir nämlich

$$y = z + 2\varepsilon \quad \text{und} \quad \psi_\nu^\varepsilon(z) = \varphi_\nu(y) \qquad \text{für } 1 \leqq \nu \leqq n,$$

so genügen die Polynome $\varphi_\nu(y)$ (in beiden Fällen) der Rekursionsgleichung

$$\varphi_{\nu+1}(y) = y\,\varphi_\nu(y) - \varphi_{\nu-1}(y) \qquad \text{für } 1 \leqq \nu < n$$

mit den Anfangsgliedern

$$\varphi_0(y) = 1; \qquad \varphi_1(y) = y.$$

Nun besteht die goniometrische Identität

$$\sin(\nu+2)\,\vartheta + \sin\nu\,\vartheta = 2\cos\vartheta\sin(\nu+1)\,\vartheta \qquad \text{für } 1 \leqq \nu < n;$$

setzen wir demgemäß

$$y = 2\cos\vartheta \quad \text{und} \quad \frac{\sin(\nu+1)\,\vartheta}{\sin\vartheta} = u_\nu(2\cos\vartheta) = u_\nu(y) \qquad \text{für } 1 \leqq \nu < n,$$

so genügen die Polynome $u_\nu(y)$ der Rekursionsgleichung

$$u_{\nu+1}(y) = y\,u_\nu(y) - u_{\nu-1}(y) \qquad \text{für } 1 \leqq \nu < n$$

mit den Anfangsgliedern

$$u_0(y) = 1; \qquad u_1(y) = y.$$

Mithin besteht die Identität

$$u_\nu(y) = \varphi_\nu(y) \qquad \text{für } 1 \leqq \nu \leqq n.$$

Infolgedessen gewinnt man die Nullstellen $\eta_1, \eta_2, \ldots, \eta_n$ des Polynoms $\varphi_n(y)$ durch die von Null verschiedenen Lösungen der Gleichung

$$\sin(n+1)\,\vartheta = 0.$$

Man findet

$$\eta_\nu = 2\cos\frac{\nu\pi}{n+1} \qquad \text{für } 1 \leqq \nu \leqq n.$$

Daher besitzen die Ausgangspolynome $\psi_n^\varepsilon(z)$ die Nullstellen

$$\zeta_\nu^\varepsilon = 2\left(\cos\frac{\nu\pi}{n+1} - \varepsilon\right) \qquad \text{für } 1 \leqq \nu \leqq n,$$

also den beiden Werten $\varepsilon = \pm 1$ entsprechend die Nullstellen

$$\zeta_\nu^+ = 2\left(\cos\frac{\nu\pi}{n+1} - 1\right) = -4\sin^2\frac{\nu}{n+1}\cdot\frac{\pi}{2} \qquad \text{für } 1 \leqq \nu \leqq n,$$

$$\zeta_\nu^- = 2\left(\cos\frac{\nu\pi}{n+1} + 1\right) = 4\sin^2\frac{n+1-\nu}{n+1}\cdot\frac{\pi}{2} \qquad \text{für } 1 \leqq \nu \leqq n.$$

In gleicher Weise erhalten wir zu der besonderen Serie

$$b_\nu = a_\nu = (-1)^\nu \qquad \text{für } 0 \leqq \nu \leqq n$$

ganzer Zahlen die Polynomserie

$$\psi_{\nu+1}(z) = z\,\psi_\nu(z) + \psi_{\nu-1}(z) \qquad \text{für } 1 \leqq \nu < n$$

mit den Anfangsgliedern

$$\psi_0(z) = 1; \qquad \psi_1(z) = z.$$

Durch Übergang zu den Polynomen

$$\varphi_\nu(y) = (-i)^\nu \psi_\nu(iy) \qquad \text{für } 0 \leqq \nu \leqq n$$

erhalten wir leicht die Rekursionsgleichung

$$\varphi_{\nu+1}(y) = y\,\varphi_\nu(y) - \varphi_{\nu-1}(y) \qquad \text{für } 1 \leqq \nu \leqq n$$

mit den Anfangsgliedern

$$\varphi_0(y) = 1; \qquad \varphi_1(y) = y.$$

Da die Nullstellen des letzten Polynoms $\varphi_n(y)$ nach dem soeben Gezeigten die algebraischen Zahlen

$$\eta_\nu = 2\cos\frac{\nu\pi}{n+1} \qquad \text{für } 1 \leqq \nu \leqq n$$

sind, erhalten wir für das Ausgangspolynom $\psi_n(z)$ die Nullstellen

$$\zeta_\nu = -2i\cos\frac{\nu\pi}{n+1} \qquad \text{für } 1 \leqq \nu \leqq n.$$

Zusammenfassend haben wir hiernach festgestellt:

Satz 5. *Für jede rationale Zahl r des Intervalls $0 < r < 1$ sind die algebraischen Zahlen*

$$\pm 4\sin^2 r\frac{\pi}{2} \qquad \textit{und} \qquad \pm 2i\cos r\frac{\pi}{2}$$

sämtlich Ausnahmewerte.

Dem offenen Kreisstreifen $\mathfrak{C}_\omega$ gehören insbesondere die offene Strecke $(-4, +4)$ der reellen Zahlengeraden und die offene Strecke $(-2i, +2i)$ der imaginären Zahlengeraden an; nach Satz 5 liegen auf diesen beiden Strecken des Kreisstreifens $\mathfrak{C}_\omega$ die Ausnahmewerte überall dicht.

Erlangen, Math. Institut

(Eingegangen am 10. Februar 1959)

THOMPSON, J. G.
Math. Zeitschr. 72, 332—354 (1960)

Normal p-complements for Finite Groups*

By
JOHN G. THOMPSON

I. Introduction

The object of this paper is to prove a theorem dealing with the p-Sylow subgroups of finite groups, for p an odd prime.

THEOREM A. *Let G be a finite group with a p-Sylow subgroup P, p an odd prime, and let $\mathfrak{A}$ be a group of automorphisms of G which leaves P invariant. Suppose for every $\mathfrak{A}$-invariant normal subgroup Q of P, elements of order prime to p which normalize Q also centralize Q. Then G possesses a normal p-complement.*

Two consequences of Theorem A are contained in [6]. No attempt will be made here to indicate the scope of Theorem A. However, Theorem A is false for the prime 2.

I would like to acknowledge my indebtedness to my thesis advisor, Professor SAUNDERS MACLANE, who gave encouragement and criticism at many points. Professor MARSHALL HALL suggested the topic and stimulated my interest in finite group theory. Professor REINHOLD BAER gave helpful advice during the summer of 1958 at the University of Frankfurt, where a large portion of the thesis was written. Special thanks are due Dr. BERTRAM HUPPERT, whose careful reading of the thesis led to a number of improvements.

The idea of the proof derives largely from HALL and HIGMAN [5] and WIELANDT [7].

II. Notation, Definitions, and Summary of Known Results

2.1. NOTATION. The notations in this paper, most of them familiar, will either be explained in context or in the list which follows.

G, H, X, Y	groups. Other letters will be used for groups, too.
$o(X)$	order of X.
$[X:Y]$	index of Y in X.
$N_X(Y)$	normalizer of Y in X.
$C_X(Y)$	centralizer of Y in X.
$Z(X)$	center of X.
$\{\ldots\|\ldots\}$	group generated by all ... such that ....
$[x, y]$	$x^{-1}y^{-1}xy = x^{-1}x^y$.
$[x, y, z]$	$[[x, y], z]$.

*) This paper was written in partial fulfillment of the requirements of a PhD at the University of Chicago.

$[X, Y]$	$\{[x, y] \mid x \in X, y \in Y\}$.
X'	$[X, X]$.
$[\ldots\|\ldots]$	the set of all ... such that ....
$\alpha_p(Y)$	the exponent of the highest power of p dividing $o(Y)$.
π	a set of primes, determined from context.
π'	the complementary set of primes to π.
π-element	an element x such that $o(x)$ involves only primes in π.
π-group	a group X such that x is a π-element for all x in X.
X_p	the smallest normal subgroup of X such that X/X_p is a p-group.
$\varphi(X)$	the intersection of all maximal subgroups of X, called the Frattini subgroup of X.
$A < X$	A is a proper subgroup of X.
$A \lhd X$	A is a normal subgroup of X. $A = X$ is permitted.
$c_q(X)$	the class of nilpotency of a q-Sylow subgroup of X.
$ccl_X(M)$	$[M^x \mid x \in X]$, M any subset of X. In case there is no danger of confusion, we put $ccl_X(M) = ccl(M)$.
$\mathrm{Aut}(X)$	the group of all automorphisms of X.
$I(X)$	the group of all inner automorphisms of X.

All groups considered are finite.

Remark. If W is an elementary abelian q-subgroup of X and $a \in X$ normalizes W, then we will consider W as a vector space over integers modulo q and the map $w \to w^a$ as a linear transformation of W. When there is no danger of confusion we will pass back and forth from the additive notation of linear transformation theory to the multiplicative notation of inner automorphisms and commutators in X.

2.2. Definitions.

Hall subgroup	H is a Hall subgroup of the group X if $(o(H), [X:H]) = 1$.
weak closure, weakly closed	If A and H are subgroups of X, then $V(ccl_X(A), H) = \{A^x \mid A^x \leqq H, x \in X\}$ is called the weak closure of $ccl_X(A)$ in H. If $A \leqq H$, and $V(ccl_X(A), H) = A$, we say that A is weakly closed in H. In case there is no danger of confusion, we put $V(H) = V(ccl_X(A), H)$ [3, p. 205].
p-normal	X is p-normal means that for any p-Sylow subgroup P of X, $Z(P)$ is weakly closed in P [3, p. 205].
π-solvable	X is π-solvable means every composition factor of X is either a π-group or a π'-group.
π-closed	X is π-closed means X possesses a normal π-subgroup K such that X/K is a π'-group. In the special case that $\pi' = [p]$, we will say that X has a normal p-complement.
$A_X(H)$	The mapping $\alpha : N_X(H) \to \mathrm{Aut}(H)$ defined by $(\alpha(x))(h) = x^{-1}hx$ has $C_X(H)$ for its kernel. We define $A_X(H)$ to be the image of $N_X(H)$ under α. $I(H) \leqq A_X(H)$ and the group $A_X(H)/I(H)$ gives one measure of the extent to which H is non-trivially imbedded in X. Note that $A_X(H) \cong N_X(H)/C_X(H)$.

2.3. SUMMARY OF KNOWN RESULTS.

a. Theorem of GRÜN-HALL. If X is p-normal, then the greatest factor group of X which is a p-group is isomorphic to that for the normalizer of the center of a p-Sylow subgroup of X. In notation, $X/X_p \cong N/N_p$, $N = N_X(Z(P))$, P a p-Sylow subgroup of X [3, p. 216].

b. If X is a p-group, then $X/\varphi(X)$ is elementary abelian [3, p. 176].

c. If σ is an automorphism of the p-group X such that σ induces the identity on $X/\varphi(X)$ and σ has order prime to p, then $\sigma = 1$ [3, p. 178].

d. If X is a p-group, then $\varphi(X)$ is generated by X' and the p-th powers of elements of X [4, p. 35].

e. First Sylow theorem. If X is of order $p^m s$, $p \nmid s$, $m \geqq 1$, p a prime, then X contains subgroups of order p^i, $i = 1, \ldots, m-1$, and each subgroup of order p^i is contained as a normal subgroup of at least one subgroup of order p^{i+1} [3, p. 44].

f. The Frattini argument. If $H \lhd X$, P a p-Sylow subgroup of H, and $N = N_X(P)$, then $X = HN$ [1, p. 117].

g. If H is a normal Hall subgroup of X, then X possesses a subgroup K such that $H \cap K = 1$, $HK = X$ [3, p. 224].

h. If X is a π-solvable group with no non-trivial normal π'-subgroups, and H is the maximal normal π-subgroup of X, then $C_X(H) = Z(H)$ [5, p. 5].

j. If A is a maximal abelian normal subgroup of the p-group X, then $A = C_X(A)$ [8, p. 145].

k. If A is a normal subgroup of order p contained in the p-group X, then A is contained in the center of X [3, p. 49].

l. Let P_1 be a p-Sylow subgroup of X, and Q_1 a characteristic subgroup of P_1. If Q_1 is not weakly closed in P_1, then there is another p-Sylow subgroup P_2 of X which contains Q_1 but in which Q_1 is not normal [3, p. 213].

m. Let P_1 be a p-Sylow subgroup of X, and H any normal subgroup of P_1 which is contained in some other p-Sylow subgroup P_2 of X as a non-normal subgroup. Then P_1 possesses a subgroup $K \geqq H$ such that $A_X(K)$ is not a p-group [3, p. 46].

THEOREM B. Let H be a p-soluble linear group over a field of characteristic p, with no normal p-subgroup greater than 1. If g is an element of order p^m in H, then the minimal equation of g is $(x-1)^r = 0$, where $r = p^m$, unless there is an integer m_0, not greater than m, such that $p^{m_0} - 1$ is a power of a prime q for which $c_q(H) > 1$, in which case, if m_0 is the least such integer,

$$p^{m-m_0}(p^{m_0} - 1) \leqq r \leqq p^m.$$

The above Theorem B of HALL and HIGMAN [5, p. 3] has been stated in full since it is so important for this paper. However, Theorem B will be applied only to groups H all of whose q-Sylow subgroups for $q \neq p$ are abelian, so the proviso of Theorem B will not be operative. Otherwise the prime $p = 3$ would cause difficulty.

III. Preliminary Lemmas

This section presents several lemmas necessary in the proof which are still of a rather general nature and make no specific reference to Theorem A.

LEMMA 3.1. *Subgroups and quotient groups of π-closed groups are π-closed.*

PROOF. Let X be π-closed and K a normal π-subgroup of X such that X/K is a π'-group. K is characterized as the set of all π-elements of X, so that π-closed groups are characterized by the property that the set of all π-elements forms a subgroup. This characterization is obviously inherited by subgroups. As for quotient groups, if X/N is any quotient group of X, then $KN/N \cong K/K \cap N$ is a normal π-subgroup of K/N, and $X/N/KN/N \cong X/KN$ is a π'-group, since X/KN is a quotient group of the π'-group X/K.

LEMMA 3.2. *If X has a normal p-complement, then X is p-normal. If P is a p-Sylow subgroup of X and $H \lhd P$, then H is normal in every p-Sylow subgroup of X which contains it.*

PROOF. Let L be the normal p-complement for P, so that $X = PL$, $P \cap L = 1$ and let K be any subgroup of P. Suppose $g \in N_X(K)$ is of order prime to p, so that $g \subset L$. Then $[k, g] \in K \cap L = 1$ for all $k \in K$, so that g centralizes K. That is, $A_X(K)$ is a p-group. It follows now from 2.3. m that if $H \lhd P$, then H is normal in every p-Sylow subgroup of X which contains it. Taking $H = Z(P)$, it follows from 2.3. l that X is p-normal.

LEMMA 3.3. *If $H \lhd X$, and P is a p-Sylow subgroup of H, then $N_X(P)$ contains a p-Sylow subgroup of X.*

PROOF. Since H is normal in X, and any two p-Sylow subgroups of H are conjugate in H, we see that $[X : N_X(P)] = [H : N_H(P)]$. Since $[H : N_H(P)]$ is prime to p, so is $[X : N_X(P)]$. In other words, $N_X(P)$ contains a p-Sylow subgroup of X.

LEMMA 3.4. *If X is a p-group and $A \neq 1$ is a normal subgroup of X, then*

i. A possesses a subgroup B of index p in A, with $B \lhd X$.

ii. $A \cap Z(X) \neq 1$.

iii. $[X, A]$ is a proper subgroup of A.

PROOF. The total number of subgroups of index p contained in A is $\equiv 1 \pmod{p}$ by a result of FROBENIUS [2, p. 152]. X induces automorphisms of A by conjugation. The subgroups of index p in A are split up into sets of transitivity under X and the number of elements in each set is a power of p. Since the total number of subgroups of index p in A is $\equiv 1 \pmod{p}$, at least one set of transitivity must contain precisely one element, that is, at least one subgroup B of index p in A must be normal in A. This proves i).

Iterating i), we find a subgroup of A of order p which is normal in X. By 2.3.k, this subgroup is in the center of X, which proves ii).

If B is a subgroup of index p in A which is normal in X, then again by 2.3.k, A/B is in the center of X/B, so that $[X, A] \leqq B$, which proves iii).

LEMMA 3.5. *Let V be a finite-dimensional vector space over a field of characteristic p, $p \neq 0$, and let X be a p-group of linear transformations of V for*

the same prime p. Let V_1 and V_2 be two X-invariant subspaces, $V_1 \supseteq V_2$ such that $V_1(I-A) \subseteq V_2$ for every element A of X. Let W_1 be the subspace of V_2 spanned by all vectors w_1 of the form $w_1 = v(I-C)$, where v is an arbitrary vector in V_1 and C is an arbitrary element of the commutator group of X. Let W_2 be the subspace of V_2 spanned by all w_2 of the form $w_2 = v(I-A)$ where v is an arbitrary vector in V_2 and A is an arbitrary element in X. Then $W_1 \subseteq W_2$.

PROOF. Suppose C has the special form $C = A^{-1}B^{-1}AB$. Every element in X has the form $I +$ nilpotent, since X is a p-group over a field of characteristic p. In particular, $A = I+N$, $B = I+M$, $C = I+N'$, and $A^{-1} = I-N+N^2\ldots$, $B^{-1} = I-M+M^2\ldots$. The relation $C = A^{-1}B^{-1}AB$ shows that N' is a polynomial in the (not necessarily commuting) variables N and M, and that each monomial is of degree at least two. Since $I-A = -N$, $I-B = -M$, and $I-C = -N'$, we find that $w_1(I-C) = -w_1N'$, and if we expand out N' as a polynomial in N and M, we see that $-w_1N'$ is a linear combination of vectors $t(I-A)$ and $t(I-B)$, where t lies in V_2. This is so since N' is a polynomial in $I-A$ and $I-B$ such that every monomial is at least of degree two, and since $w_1(I-A)$ and $w_1(I-B)$ both lie in V_2 by assumption. In other words, we have proved the lemma in the special case that $C = A^{-1}B^{-1}AB$.

Now let C be any element in the commutator group of X. Then $C = C_1C_2\ldots C_r$, where $C_i = A_i^{-1}B_i^{-1}A_iB_i$ for suitable A_i and B_i. If now we make use of the identity

$$I - C = I - C_1 + C_1(I-C_2) + C_1C_2(I-C_3) + \cdots + C_1\ldots C_{r-1}(I-C_r),$$

and also make use of the fact that both V_1 and V_2 are X-invariant, the entire lemma follows.

LEMMA 3.6. *Let X be a p-group and $\mathfrak{B}$ a group of automorphisms of X. Suppose A and B are two $\mathfrak{B}$-invariant normal subgroups of X and $A < B$. Then we can find a chain of $\mathfrak{B}$-invariant normal subgroups of X, $A = A_1 < A_2 < \cdots < A_n = B$, such that $[X, A_{i+1}] \subseteq A_i$, $i = 1, \ldots, n-1$.*

PROOF. A_{k+1} can be taken as the intersection of B with the complete inverse image of $Z(X/A_k)$. A_{k+1} will be $\mathfrak{B}$-invariant and normal if A_k is, and since $A_1 = A$ is $\mathfrak{B}$-invariant and normal, we see that A_k is also $\mathfrak{B}$-invariant and normal. Since $B/A_k \lhd X/A_k$, it follows from lemma 3.4.ii that $Z(X/A_k) \cap B/A_k \neq 1$, provided only that A_k is a proper subgroup of B, so that A_{k+1} contains A_k as a proper subgroup. We repeat this process until we reach B.

Let $\mathfrak{B}$ be a group of automorphisms of the p-group X. We say that a subgroup A of X is of type I relative to $\mathfrak{B}$ if and only if

1. A is $\mathfrak{B}$-invariant and normal in X, and $A \neq 1$.
2. $[X, A, A] = 1$.

Note that all abelian $\mathfrak{B}$-invariant normal subgroups of X ($\neq 1$) are of type I relative to $\mathfrak{B}$.

We say that the subgroup A of X is of type II relative to $\mathfrak{B}$ if and only if

1. A is non-abelian of class 2.
2. $[X, A] \leqq Z(A)$.
3. $A/Z(A)$ is elementary abelian.
4. A is $\mathfrak{B}$-invariant and normal in X.
5. $Z(X) \leqq A$.

Observe that subgroups of type II are also subgroups of type I relative to $\mathfrak{B}$.

LEMMA 3.7. *Let X be a p-group and $\mathfrak{B}$ a group of automorphisms of X. Suppose that none of the maximal abelian normal subgroups of X is $\mathfrak{B}$-invariant. Then X possesses a subgroup of type II relative to $\mathfrak{B}$.*

PROOF. There are certainly $\mathfrak{B}$-invariant abelian normal subgroups of X which contain $Z(X)$, since $Z(X)$ is such a subgroup. Choose B maximal among all such. It cannot occur that $B = C(B)$, since that would imply that B was a maximal abelian normal subgroup of X which was also $\mathfrak{B}$-invariant. Hence $B < C(B)$. In $C(B)$, let us consider those subgroups C such that

1. C is $\mathfrak{B}$-invariant and normal in X.
2. C is not contained in B.

Since $C(B)$ satisfies these two conditions, it follows that $C(B)$ possesses a subgroup with these two properties and is of minimal order, say C. Then $\varphi(C)$ must be contained in B, since $\varphi(C)$ is $\mathfrak{B}$-invariant and normal along with C, and is of smaller order than C. Also $[X, C]$ must be contained in B, since it too is $\mathfrak{B}$-invariant and normal along with C, and by lemma 3.4.iii is of smaller order than C. Put $A = BC$. Remark that $B \leqq Z(A)$, since B is abelian and C centralizes B. Now $[X, BC] = [X, B][X, C]$, since $[x, bc] = [x, c][x, b]^c$, and B and C are both normal in X. Since $[X, B]$ and $[X, C]$ are both contained in B, we find $[X, A] \leqq B \leqq Z(A)$. Thus, A has properties 2, 4, and 5 in the definition of type II.

Now A is definitely not abelian, since A is $\mathfrak{B}$-invariant and normal and contains B as a proper subgroup, and B has been chosen as a maximal abelian $\mathfrak{B}$-invariant normal subgroup. On the other hand, $[A, A] \leqq [X, A] \leqq Z(A)$, which shows that A has class at most two. Hence, A has class exactly two, that is, A has property 1 in the definition of type II.

Since $A/\varphi(A)$ is elementary abelian by 2.3.b, we shall show that $A/Z(A)$ is elementary abelian by showing that $\varphi(A) \leqq Z(A)$. Now $\varphi(A)$ is generated by the commutator group of A and by the p-th powers of elements of A, by 2.3.d. Since $[A, A] \leqq Z(A)$, it suffices to prove that p-th powers of elements of A lie in $Z(A)$. Since $B \leqq Z(A)$, we have $(bc)^p = b^p c^p$. Since $b^p \in B \leqq Z(A)$, we must show that $c^p \in Z(A)$. But $c^p \in \varphi(C) \leqq B \leqq Z(A)$. This completes the proof of this lemma.

LEMMA 3.8. *Let $A \leqq B \leqq X$, where A and B are normal subgroups of the p-group X, and where $[X, B] \leqq A$. Then B normalizes every subgroup of X which contains A.*

PROOF. Since $[X, B] \leqq A$, we see that $B/A \leqq Z(X/A)$, so that B/A normalizes every subgroup of X/A. It follows that B normalizes every subgroup of X which contains A.

LEMMA 3.9. *Let X be a p-normal group, P a p-Sylow subgroup of X, and A a subgroup of $Z(P)$. Suppose that A is contained in the p-Sylow subgroup P^* of X. Then $A \leqq Z(P^*)$.*

PROOF. Put $C = C_X(A)$. We will prove the lemma by induction on $o(X)$. If $C = X$, then A is in $Z(X)$, and the lemma is true. If $C < X$, then the group $B = \{A, Z(P^*)\}$ is contained in some p-Sylow subgroup P_1 of C, and since P is contained in C, P_1 is even a p-Sylow subgroup of X. Since the p-Sylow subgroups of C are also p-Sylow subgroups of X, it follows that C is p-normal. Since $o(C) < o(X)$, we see that $A \leqq Z(P_1)$. But $Z(P^*) \leqq P_1$, and since X is p-normal, we must have $Z(P_1) = Z(P^*)$. Hence $A \leqq Z(P^*)$, as desired.

LEMMA 3.10. *If X possesses a normal p-complement, $o(X) = \prod_q q^{\alpha_q(X)}$, then X possesses a subgroup of order $p^{\alpha_p} q^{\alpha_q}$ in which the q-Sylow subgroup is normal. Here q is any prime divisor of $o(X)$ different from p.*

PROOF. Let K be the normal complement to a p-Sylow subgroup of X, and let Q be a q-Sylow subgroup of K, $q \mid o(K)$. Since K is normal in X, $[K : N_K(Q)] = [X : N_X(Q)]$. Since $o(K)$ is prime to p, $[K : N_K(Q)]$ is also prime to p. Hence $[X : N_X(Q)]$ is prime to p, that is, $N_X(Q)$ contains a p-Sylow subgroup of X, say P. The group PQ satisfies the demands of the lemma.

LEMMA 3.11. *If X is a group of linear transformations on a finite dimensional vector space V, and X acts irreducibly on V, then $Z(X)$ is cyclic.*

PROOF. By a theorem of SCHUR, the set of linear transformations of V which centralize X forms a skew field. $Z(X)$ is imbedded in this skew field as a finite subgroup of the multiplicative subgroup. But finite abelian subgroups of the multiplicative group of a field are cyclic.

IV. Initial Reduction of the Proof of Theorem A

Theorem A is a statement dealing with triples $(X, Y, \mathfrak{B})$ where

1. Y is a p-Sylow subgroup of X for an odd prime p.
2. $\mathfrak{B}$ is a group of automorphisms of X such that $Y^{\mathfrak{B}} = Y$.
3. If H is any $\mathfrak{B}$-invariant normal subgroup of Y, then $A_X(H)$ is a p-group.

Triples $(X, Y, \mathfrak{B})$ satisfying these three conditions, that is, satisfying the hypotheses of Theorem A, will be said to have property $\mathfrak{N}$. This serves to define property $\mathfrak{N}$, and is a convenient device for future reference.

We can now state Theorem A in a concise way.

THEOREM A. *Let $(G, P, \mathfrak{A})$ have property $\mathfrak{N}$. Then G has a normal p-complement.*

LEMMA 4.1. *Let $(G, P, \mathfrak{A})$ have property $\mathfrak{N}$. Suppose that Theorem A has been proved for all triples $(G_1, P_1, \mathfrak{A}_1)$ with property $\mathfrak{N}$ with $o(G_1) < o(G)$. Suppose further that G is p-normal. Then G has a normal p-complement, that is, Theorem A is true for the triple $(G, P, \mathfrak{A})$.*

PROOF. If G is either a p-group or a p'-group, then Theorem A is trivially true, so we can suppose that G is neither a p-group nor a p'-group.

The theorem of GRÜN-HALL (2.3.a), applied to G, assumed to be p-normal, gives $G/G_p \cong N/N_p$, where $N = N_G(Z(P))$, P a p-Sylow subgroup of G.

CASE 1. $N < G$. Since P is $\mathfrak{A}$-invariant, so is $Z(P)$, since characteristic subgroups of $\mathfrak{A}$-invariant subgroups are $\mathfrak{A}$-invariant. Since $Z(P)$ is $\mathfrak{A}$-invariant, so is N, since normalizers of $\mathfrak{A}$-invariant subgroups are $\mathfrak{A}$-invariant. Let $\mathfrak{A}_1$ be $\mathfrak{A}$ restricted to N. One checks that the triple $(N, P, \mathfrak{A}_1)$ has property $\mathfrak{N}$. Since $o(N) < o(G)$, we conclude that N possesses a normal p-complement. Since $G/G_p \cong N/N_p$, G possesses a normal p-complement, too.

CASE 2. $N = G$, that is, $Z = Z(P)$ is normal in G. Since elements in G of order prime to p centralize Z by condition 3 defining property $\mathfrak{N}$, and since P centralizes Z, it follows that $Z \leqq Z(G)$. Let $\mathfrak{A}'$ be the group of automorphisms which $\mathfrak{A}$ induces on G/Z. One checks that the triple $(G/Z, P/Z, \mathfrak{A}')$ has property $\mathfrak{N}$. Since $o(G/Z) < o(G)$, we conclude that G/Z has a normal p-complement, and it follows readily from the fact that Z is in the center of G that G itself has a normal p-complement.

Lemma 4.1 enables us to throw our attention onto the question of the p-normality of G, since Theorem A is to be proved by induction on $o(G)$.

V. Further Lemmas

We assume now by way of contradiction that Theorem A is false and choose a triple $(G, P, \mathfrak{A})$ with property $\mathfrak{N}$ which does not satisfy the conclusion of Theorem A, and such that $o(G)$ is minimal. The next few lemmas deal with this triple $(G, P, \mathfrak{A})$, although we don't yet attack the heart of the matter.

Since we are assuming that Theorem A is false for the triple $(G, P, \mathfrak{A})$, it follows from lemma 4.1 and from the minimal nature of G that G is not p-normal. In other words, the proof of Theorem A will be complete when we have obtained a contradiction from the assumption that G is not p-normal.

LEMMA 5.1. *$Z(P)$ is contained in a p-group M of G, such that $A_G(M)$ is not a p-group.*

PROOF. This follows immediately from 2.3.1 and 2.3.m and the non-p-normality of G.

LEMMA 5.2. *If P^x is another p-Sylow subgroup of G, and*

$$\mathfrak{A}^x = \{C_x \sigma C_{x^{-1}} \mid \sigma \in \mathfrak{A}\},$$

where C_x is the inner automorphism of G induced by x, then the triple $(G, P^x, \mathfrak{A}^x)$ also has property $\mathfrak{N}$.

PROOF. Let $H_1 = x^{-1} H x$ be an $\mathfrak{A}^x$-invariant normal subgroup of P^x, where H is a subgroup of P. To say that H_1 is $\mathfrak{A}^x$-invariant is to say $(C_x \sigma C_{x^{-1}})(H_1) = H_1 = x^{-1} H x$ for all $\sigma \in \mathfrak{A}$. But $(C_x \sigma C_{x^{-1}})(H_1) = x^{-1} \sigma(H) x$, so we find that $\sigma(H) = H$, for all $\sigma \in \mathfrak{A}$. Thus H is an $\mathfrak{A}$-invariant subgroup of P. If $H_1 \lhd P^x$, then $H \lhd P$, and by assumption $A_G(H)$ is a p-group. Since H and H_1 are conjugate in G, $A_G(H_1)$ is also a p-group. That is, the triple $(G, P^x, \mathfrak{A}^x)$ has property $\mathfrak{N}$.

LEMMA 5.3. *If $B \neq 1$ is any $\mathfrak{A}$-invariant normal subgroup of P, then $N_G(B)$ possesses a normal p-complement.*

PROOF. Put $N = N_G(B)$. Then $o(N/B) < o(G)$. N is $\mathfrak{A}$-invariant along with B, so let $\mathfrak{A}_1$ be $\mathfrak{A}$ restricted to N/B. The triple $(N/B, P/B, \mathfrak{A}_1)$ is seen to have property $\mathfrak{N}$, so by the induction assumption has a normal p-complement. This guarantees the existence of a normal subgroup K/B of N/B such that $o(K/B)$ is prime to p, while $o(N/K)$ is a power of p. By 2.3.g, we can find in K a subgroup L such that $L \cap B = 1$, $LB = K$. Since $o(L)$ is prime to p, and since B is $\mathfrak{A}$-invariant and normal in K, elements in L centralize B. Thus $K = L \times B$, and L is characteristic in K, hence normal in N. This shows that L is a normal p-complement for N.

LEMMA 5.4. *If $B \neq 1$ is any $\mathfrak{A}$-invariant normal subgroup of P, then B is not normal in G.*

PROOF. This is an immediate corollary of lemma 5.3, since G does not have a normal p-complement.

LEMMA 5.5. *G possesses no non-trivial normal p-groups.*

PROOF. The maximal normal p-subgroup of G is characteristic in G, hence $\mathfrak{A}$-invariant. Being normal it is contained in every p-Sylow subgroup of G, in particular in P. By lemma 5.4, this implies that the maximal normal p-subgroup of G is 1.

LEMMA 5.6. *Suppose B and C are $\mathfrak{A}$-invariant normal subgroups of P such that*

i. $1 < B \leqq C$.
ii. $[P, C] \leqq B$.
iii. $[B, C] = 1$.

Let T be any p-group contained in $C_G(B)$ such that $\{C, T\} = U$ is a p-group. Then $[T, C] \leqq B$, so that $[T, C, C] = 1$.

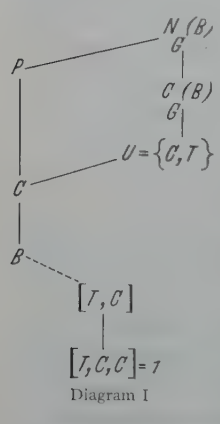


Diagram I

PROOF. Diagram I gives a picture of what we are trying to prove. By lemma 5.3, $N = N_G(B)$ has a normal p-complement. Hence, N/B has a normal p-complement. Now P/B is a p-Sylow subgroup of N/B. Since $[P, C] \leqq B$, the group C/B is contained in $Z(P/B)$. Since N/B has a normal p-complement, it is p-normal by lemma 3.2. Consider the group U/B. The lemma requires that we prove $[C, U] \leqq B$. Now U/B, being a homomorphic image of the p-group U, is a p-group, so U/B is contained in a p-Sylow subgroup P^*/B of N/B. Since N/B is p-normal, we conclude from lemma 3.9 that $C/B \leqq Z(P^*/B)$, and the lemma follows.

LEMMA 5.7. *Suppose A is a maximal abelian normal subgroup of P. Then any p-element of G which centralizes A lies in A.*

PROOF. Put $N = N_G(A)$ and $C = C_G(A)$. C is a normal subgroup of N, and the lemma demands that we show that A is a p-Sylow subgroup of C. Let P_1 be a p-Sylow subgroup of C. By lemma 3.3, the normalizer of P_1 in N contains a

p-Sylow subgroup of N, say P_2. Since P is a p-Sylow subgroup of N, we have $P_2 = P^t$ for some $t \in N$. But then P normalizes $P_1^{t^{-1}}$, so we see that P normalizes some p-Sylow subgroup of C. We then have the relations $A \leqq P_1^{t^{-1}} \leqq P$. Since $P_1^{t^{-1}}$ centralizes A, we conclude by 2.3.j that $A = P_1^{t^{-1}}$, so that A is a p-Sylow subgroup of $C_G(A)$. Since $A \lhd C_G(A)$, A is the p-Sylow subgroup of C, and this completes the proof of the lemma.

LEMMA 5.8. *If A is any $\mathfrak{A}$-invariant normal subgroup of P, then $V = V(ccl_G(A), P)$ is an $\mathfrak{A}$-invariant normal subgroup of P.*

PROOF. Let A^x be any conjugate of A contained in P. Then for $\sigma \in \mathfrak{A}$, we have $\sigma(A^x) = \sigma(x)^{-1} A \sigma(x)$, since A is $\mathfrak{A}$-invariant. Thus σ sends conjugates of A into conjugates of A. Since σ leaves P invariant, it merely permutes the conjugates of A which lie in P among themselves and leaves V invariant. Since this is true for all $\sigma \in \mathfrak{A}$, we conclude that V is $\mathfrak{A}$-invariant. It is obviously normal in P.

LEMMA 5.9. *If $A \neq 1$ is any $\mathfrak{A}$-invariant normal subgroup of P, $A \leqq D \leqq P$, and R is a q-subgroup of G which normalizes D but does not centralize D, q a prime different from p, then R does not centralize A.*

PROOF. Suppose false and that $R \leqq C_G(A)$. By lemma 5.3, $N_G(A)$ has a normal p-complement. By assumption, R is contained in $N_G(A)$ (it is even contained in $C_G(A)$), and R normalizes D, which is also contained in $N_G(A)$, since P is contained in $N_G(A)$. The group $\{R, D\} = RD$, being a subgroup of $N_G(A)$, has a normal p-complement by lemma 3.1, which is necessarily R. On the other hand, R normalizes D by assumption. Hence R and D, being normal subgroups of RD of coprime orders, commute elementwise, that is, R centralizes D. This contradiction completes the proof of lemma 5.9.

VI. Subgroups of Types I, II, and III

We now define a partial ordering $\ll$ on the set of all subgroups of G. If H_1 and H_2 are subgroups of G, we say that $H_1 \ll H_2$ if and only if $\alpha_p(N_G(H_1)) < \alpha_p(N_G(H_2))$, or $\alpha_p(N_G(H_1)) = \alpha_p(N_G(H_2))$ and $o(H_1) < o(H_2)$, or $H_1 = H_2$. Note that conjugate subgroups are not comparable. If $\mathfrak{S}$ is any set of subgroups of G, and $H \in \mathfrak{S}$, we say that H is maximal in the ordering $\ll$, relative to $\mathfrak{S}$, provided that whenever $K \in \mathfrak{S}$ and $H \ll K$, then $H = K$.

If A is any $\mathfrak{A}$-invariant normal subgroup of P, we define

$$\mathfrak{H}(ccl_G(A)) = [H \mid 1.\ H \text{ is a } p\text{-subgroup of } G.$$
$$2.\ A_G(H) \text{ is not a } p\text{-group.}$$
$$3.\ H \text{ contains some element of } ccl(A).].$$

It may occur that $\mathfrak{H}(ccl_G(A))$ is empty for some choices of A. However, lemma 5.1 says that $\mathfrak{H}(ccl_G(Z(P)))$ is not empty, so we restrict ourselves to those $\mathfrak{A}$-invariant normal subgroups $A \neq 1$ of P such that $\mathfrak{H}(ccl(A))$ is not empty. Lemmas 6.1 through 6.9 all involve the following objects:

1. An $\mathfrak{A}$-invariant normal subgroup $A \neq 1$ of P with $\mathfrak{H}(ccl(A))$ not empty.
(*) 2. A subgroup $H \in \mathfrak{H}(ccl(A))$ maximal in the ordering $\ll$ relative to $\mathfrak{H}(ccl(A))$.
3. A p-Sylow subgroup P_1 of $N = N_G(H)$.

LEMMA 6.1. *If $H<J\lhd P_1$, then $A_G(J)$ is a p-group.*

PROOF. Otherwise, $J\in\mathfrak{H}(ccl(A))$ and H would not be maximal in the ordering $\ll$ relative to $\mathfrak{H}(ccl(A))$, either by virtue of the fact that $\alpha_p(N)<\alpha_p(N_G(J))$ (we always have $\alpha_p(N)\leqq\alpha_p(N_G(J))$ since P_1 is a p-Sylow subgroup of N) or because $\alpha_p(N)=\alpha_p(N_G(J))$, in which case we have $o(H)<o(J)$.

LEMMA 6.2. *$N_G(H)/H$ has a normal p-complement.*

PROOF. Let $\mathfrak{A}_1$ be the group of automorphisms of N/H induced by the elements of P_1/H by conjugation. Suppose $H<K\lhd P_1$. By lemma 6.1, $A_G(K)$ is a p-group. In particular, $A_N(K)$ is a p-group, from which it follows that $A_{N/H}(K/H)$ is a p-group, too. This guarantees that the triple $(N/H, P_1/H, \mathfrak{A}_1)$ has property $\mathfrak{N}$. Since $o(N/H)<o(G)$, it follows from the minimal nature of G that N/H has a normal p-complement.

LEMMA 6.3. *If $o(N)=\prod_q q^{\alpha_q}$, and $o(H)=p^n$, then N/H possesses a subgroup of order $p^{\alpha_p-n}q^{\alpha_q}$ in which the q-Sylow subgroup is normal, q being any prime divisor of $o(N)$ different from p.*

PROOF. By lemma 6.2, N/H has a normal p-complement. By lemma 3.10 N/H possesses a subgroup satisfying the demands of the lemma.

LEMMA 6.4. *N possesses a subgroup of order $p^{\alpha_p}q^{\alpha_q}$ for any prime $q\mid o(N)$, $q\neq p$.*

PROOF. This is an immediate consequence of lemma 6.3, since the complete inverse image in N of a subgroup of N/H satisfying lemma 6.3 satisfies lemma 6.4.

REMARK. The object of the preceding lemmas has been to produce certain solvable groups contained in G, and which involve the prime p to a "considerable" amount. The solvable groups which lemma 6.4 gives us will next be subjected to a detailed study.

LEMMA 6.5. *N possesses a subgroup of order $p^{\alpha_p}q^{\alpha_q}$, for any prime $q\mid o(A_G(H))$, $q\neq p$, in which the q-Sylow subgroups are not normal.*

PROOF. Lemma 6.4 guarantees the existence of a subgroup of N of order $p^{\alpha_p}q^{\alpha_q}$, say $K(q)$. Since a p-Sylow subgroup of $K(q)$ is also a p-Sylow subgroup of N, and since H is contained in every p-Sylow subgroup of N, H is contained in $K(q)$. Since $K(q)$ is a subgroup of N, H is normal in $K(q)$. If a q-Sylow subgroup of $K(q)$ were normal in $K(q)$, then it would commute elementwise with H. But since a q-Sylow subgroup of $K(q)$ is also a q-Sylow subgroup of N, this would say that $o(A_G(H))$ was prime to q, contrary to our choice of q.

REMARK. Since $H\in\mathfrak{H}(ccl(A))$, there is at least one prime $q\neq p$ which divides $o(A_G(H))$. We choose one once and for all, say q. The rest of the proof of Theorem A is devoted to a study of subgroups of N which involve the primes p and q only. From this point onward, p will be the prime occurring in the statement of Theorem A, and q will be an arbitrary but fixed prime divisor of $o(A_G(H))$, $q\neq p$, and no further reference will be made to this fact.

LEMMA 6.6. If K is a subgroup of N, $o(K) = p^{\alpha_p(N)} q^b$, if the q-Sylow subgroups of K are not normal in K, and if Q_1 is the maximal normal q-subgroup of K, then HQ_1/Q_1 is the maximal normal p-subgroup of K/Q_1.

PROOF. Let L be the maximal normal subgroup of K which contains Q_1 and such that L/Q_1 is a p-group, and let H_1 be a p-Sylow subgroup of L, so that $L = Q_1 H_1$. By 2.3.f, $L \cdot N_K(H_1) = K$. Hence

$$(**) \qquad (Q_1 H_1) N_K(H_1) = Q_1 N_K(H_1) = K.$$

By lemma 3.3, $N_K(H_1)$ contains a p-Sylow subgroup of K. Hence $\alpha_p(N_G(H_1)) \geqq \alpha_p(N)$, since a p-Sylow subgroup of K is also a p-Sylow subgroup of N. Also, $H \leqq H_1$, since HQ_1 is normal in K. Suppose by way of contradiction that $H < H_1$. Since H is maximal in the ordering $\ll$ relative to $\mathfrak{H}(ccl(A))$, H_1 is not contained in $\mathfrak{H}(ccl(A))$, which implies that $A_G(H_1)$ is a p-group. In particular, a q-Sylow subgroup of $N_K(H_1)$ must centralize H_1. By (**), a q-Sylow subgroup of K/Q_1 must centralize $H_1 Q_1/Q_1 = L/Q_1$. But K/Q_1 has no non-trivial normal q-subgroups, so that L/Q_1 must contain its centralizer in K/Q_1, by 2.3.h. Thus, a q-Sylow subgroup of K/Q_1 must be the identity, that is, Q_1 is a q-Sylow subgroup of K, contrary to the assumption that the q-Sylow subgroups of K are not normal in K. Hence, $L/Q_1 = HQ_1/Q_1$ is the maximal normal p-subgroup of K/Q_1.

LEMMA 6.7. *If K and Q_1 are as in lemma* 6.6, *then a q-element $x \in K$ centralizes H if and only if $x \in Q_1$.*

PROOF. By the definition of Q_1, K/Q_1 has no non-trivial normal q-subgroups. HQ_1/Q_1 is the maximal normal p-subgroup of K/Q_1 by lemma 6.6. By 2.3.h, $C_{K/Q_1}(HQ_1/Q_1) = Z(HQ_1/Q_1)$. Let $x \in K$ centralize H, x being a q-element. Then xQ_1 centralizes HQ_1/Q_1, so that $xQ_1 \in Z(HQ_1/Q_1)$. But xQ_1 is a q-element of K/Q_1, and it lies in the p-group $Z(HQ_1/Q_1)$. Hence, xQ_1 is the identity element of K/Q_1, that is, x lies in Q_1. Conversely, if $x \in Q_1$, then x centralizes H since H and Q_1 are normal subgroups of K of coprime orders and so commute elementwise.

LEMMA 6.8. *If K and Q_1 are as in lemma* 6.6, *and L is a normal subgroup of K which contains HQ_1 properly and such that L/HQ_1 is a minimal normal subgroup of K/HQ_1, and P_1 is a p-Sylow subgroup of K, then the group $K_1 = P_1 L$ has the following properties:*

1. *L/HQ_1 is an elementary abelian q-group.*
2. *A q-Sylow subgroup of K_1 is not normal in K_1.*
3. *Q_1 is the maximal normal q-subgroup of K_1.*
4. *HQ_1/Q_1 is the maximal normal p-subgroup of K_1/Q_1.*

PROOF. By lemma 6.2, N/H has a normal p-complement. Hence, by lemma 3.1, K/H has a normal p-complement. Since $o(K)$ only involves the primes p and q, K/H is solvable. (Note that we are not forced to rely on BURNSIDE's theorem that groups of order $p^a q^b$ are solvable.) Since K/HQ_1 has no non-trivial normal p-subgroups by lemma 6.6 (here we use the fact

that $K/HQ_1 \cong K/Q_1/HQ_1/Q_1$) a minimal normal subgroup of K/HQ_1 is an elementary abelian q-group. This proves the first assertion of lemma 6.8.

Since H is normal in K_1, any normal q-subgroup of K_1 centralizes H. By lemma 6.7, Q_1 is the maximal normal q-subgroup of K_1. This proves both assertions 2 and 3 of lemma 6.8. Hence, K_1 satisfies the hypotheses of lemma 6.6 with K_1 playing the role of K. We conclude that HQ_1/Q_1 is the maximal normal p-subgroup of K_1/Q_1.

We are now ready to define the groups of type III, or more precisely, the groups of type III with respect to H, H being in $\mathfrak{H}(ccl(A))$ and maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$. The necessary components in the definition include (*) preceding lemma 6.1.

The subgroup K_1 of G will be said to be of type III with respect to H if and only if:

1. $K_1 \leqq N_G(H) = N$.
2. $o(K_1) = p^{\alpha_p(N)} q^b$.
3. A q-Sylow subgroup of K_1 is not normal in K_1.
4. If Q_1 is the maximal normal q-subgroup of K_1, then K_1/Q_1 has q-Sylow subgroups which are elementary abelian.
5. Q_1 is the set of all q-elements of K_1 which centralize H.
6. HQ_1/Q_1 is the maximal normal p-subgroup of K_1/Q_1.

In intuitive terms, K_1 is chosen to contain a p-Sylow subgroup of N, and to be small subject to the restriction that it not possess a normal p-complement.

We summarize lemmas 6.5 through 6.8 by using the definition of a subgroup of type III.

LEMMA 6.8′. *If $A \neq 1$ is an $\mathfrak{A}$-invariant normal subgroup of P, $\mathfrak{H}(ccl(A))$ is not empty, and if $H \in \mathfrak{H}(ccl(A))$ is maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$, then there exists a subgroup of type III with respect to H, containing any given p-Sylow subgroup of $N_G(H)$.*

The next lemmas bring into play for the first time Theorem B of HALL and HIGMAN, and in the application of Theorem B it will be seen that the assumption that p is odd will be used. We make appeal to subgroups of types I and II defined in section 3, where we take $\mathfrak{A}$ to play the role of $\mathfrak{B}$ and P to play the role of X. Subgroups of type II will be used explicitly only in section 8. However, since subgroups of type II are also of type I, they occur implicitly in all the lemmas dealing with subgroups of type I.

LEMMA 6.9. *If the p-Sylow subgroup P_1 of N is contained in the p-Sylow subgroup $\tilde{P} = P^y$ of G, and if B is any subgroup of P of type I, then $B^y \leqq H$.*

PROOF. By conjugating with y^{-1}, we can suppose that $P_1 \leqq P$, and try to prove that H contains every subgroup of type I. This conjugation by y^{-1} is permissible, since if H is maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$, so is $H^{y^{-1}}$. To carry out the proof, we make use of a subgroup K of type III with respect to H which contains P_1.

To prove lemma 6.9, it suffices to show that every subgroup of type I which contains $Z(P)$ is contained in H, since if C is any subgroup of type I, then $B = CZ(P)$ is also of type I, so that if we can show that B is contained in H it will follow that C is also contained in H.

We first show that $Z(P)$ is contained in H. Since $H \leqq P$, $Z(P)$ is at least contained in K, since it is contained in P_1. On the other hand, HQ_1/Q_1 is the maximal normal p-subgroup of K/Q_1 and K/Q_1 has no non-trivial normal q-subgroups. By 2.3.h, HQ_1/Q_1 contains its centralizer in K/Q_1. Since $Z(P)$ centralizes H and lies in K, we see that $Z(P)$ must lie in HQ_1, and hence in H, since H is the set of all p-elements of HQ_1.

Next let B be any subgroup of type I which contains $Z(P)$ as a proper subgroup. Put $Z(P) = A_1$, and use lemma 3.6 to construct a chain of subgroups of type I, terminating at B, $A_1 < A_2 \ldots < A_n = B$, such that $[P, A_{i+1}] \leqq A_i$ for $i = 1, \ldots, n-1$. We shall show that if A_k is contained in H, $k \leqq n-1$, then so is A_{k+1}.

If A_k is contained in H, then A_{k+1} is at least contained in P_1, since by lemma 3.8, A_{k+1} normalizes H. On the other hand, if in the group K above, we represent K as linear transformations on $H/\varphi(H)$, we find by 2.3.c and lemma 6.7 that Q_1 is a q-Sylow subgroup of the kernel M of this homomorphism of K into linear transformations of $H/\varphi(H)$. H is also contained in M. M/Q_1 is a normal p-subgroup of K/Q_1 which contains HQ_1/Q_1. Since HQ_1/Q_1 is the maximal normal p-subgroup of K/Q_1, we conclude that $M = HQ_1$. That is, we have a faithful representation of K/HQ_1 as linear transformations on $H/\varphi(H)$.

It is at this point that we want to make use of Theorem B of Hall and Higman. We observe that K/HQ_1 has no non-trivial normal p-subgroups, and that a q-Sylow subgroup of K/HQ_1 is elementary abelian. Application of Theorem B shows that the minimal polynomial of an element aHQ_1 of order $p^m = r$ is $(x-1)^r$. Choose $a \in A_{k+1}$, $h \in H$. Since $[P, A_{k+1}, A_{k+1}] = 1$, we find $[h, a, a] = 1$, which says that a has a minimal polynomial on $H/\varphi(H)$ which is a divisor of $(x-1)^2$. Since $p \geqq 3$, we conclude that $r = 1$ and hence that $a \in H$. Hence $A_{k+1} \leqq H$, and the proof of the lemma is completed.

Lemma 6.10. *If A is any subgroup of type I, then $\mathfrak{H}(ccl(A))$ is not empty.*

Proof. We are required to show that if A is any subgroup of type I then there is some p-subgroup H which contains a conjugate of A and such that $A_G(H)$ is not a p-group. We know by lemma 5.1 that $\mathfrak{H}(ccl(Z(P)))$ is not empty. Lemma 6.9 gives immediately that $\mathfrak{H}(ccl(A))$ is not empty, since if H is maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(Z(P)))$, then H is contained in $\mathfrak{H}(ccl(A))$.

VII. Subgroups of Type III

The previous section was devoted essentially to proving the existence of subgroups of type III with respect to H. In this section, the considerably harder task of studying the structure of such subgroups is undertaken. This study leads to an examination of quite a number of subgroups, all of which

are relevant to the discussion more or less simultaneously. Diagram II, which follows with an explanation of the groups involved, may help the reader keep the various components in mind.

The explanation of the diagram will proceed by reconstructing the lattice, starting from the empty lattice.

1. Select a subgroup A of type I.

2. Choose $H \in \mathfrak{H}(ccl(A))$ maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$, and such that $P \cap N_G(H) = P_1$ is a p-Sylow subgroup of $N_G(H)$. This can always be accomplished by taking an arbitrary group H maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$, and passing to a conjugate of H if necessary.

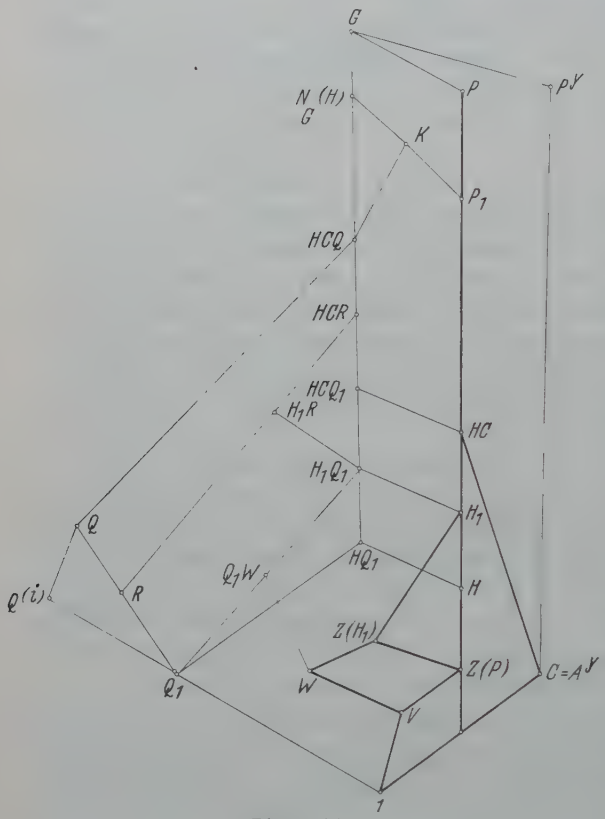


Diagram II

I. Insert $H, P_1, P, N_G(H)$, and G in the lattice.

3. Choose K to be a subgroup of type III with respect to H, and containing P_1. Denote by Q a q-Sylow subgroup of K, and by Q_1 the maximal normal q-subgroup of K.

II. Insert K, Q, and Q_1 in the lattice.

4. Select $C = A^y$ an arbitrary conjugate of A which is contained in P_1. There is at least one such C, since $H \in \mathfrak{H}(ccl(A))$.

III. Insert C and P^y in the lattice. It may occur that $P^y = P$, but this makes no difference.

5. Form the groups $\{H, C\} = HC$ (H is normal in K), and $\{H, C, Q\} = HCQ$. (Since $N_G(H)/H$ has a normal p-complement by lemma 6.2, so does K/H, that is, QH/H is normal in K/H, so that QH is normal in K.)

IV. Insert HQ_1, HC, HCQ_1, and HCQ in the lattice.

6. Q/Q_1 is elementary abelian by condition 4 in the definition of a subgroup of type III with respect to H, and $Q/Q_1 \cong HQ/HQ_1$. Also, HQ/HQ_1 is normal in HCQ/HQ_1, so that HCQ_1/HQ_1 induces automorphisms on

HQ/HQ_1. Since $HCQ_1/HQ_1 \cong C/C\cap H$, we have the p-group HCQ_1/HQ_1 represented as linear transformations on the elementary abelian q-group HQ/HQ_1. Since $p \neq q$, this representation is completely reducible. Select $Q^{(1)}, \ldots, Q^{(r)}$ such that $Q_1 < Q^{(i)} \leqq Q$, $i = 1, \ldots, r$, $r \geqq 1$, and such that $HQ^{(i)}/HQ_1$ is a minimal HCQ_1/HQ_1-invariant subgroup of HQ/HQ_1, and such that $Q/Q_1 = Q^{(1)}/Q_1 \times \cdots \times Q^{(r)}/Q_1$. The groups $Q^{(i)}$ are simply Q intersected with the complete inverse image in HCQ of minimal HCQ_1/HQ_1-invariant subgroups of HQ/HQ_1.

V. Insert the groups $Q^{(i)}$ in the lattice.

7. Choose a fixed but arbitrary index j, $1 \leqq j \leqq r$, and put $R = Q^{(j)}$.

VI. Insert R and HCR in the lattice.

8. Select $H_1 \leqq HC$, such that H_1Q_1 is normal in HCR, and such that H_1Q_1/Q_1 is the maximal normal p-group of HCR/Q_1. $H \leqq H_1$, since HQ_1 is normal in HCR, and since H_1 is the join of all subgroups T of HC such that TQ_1 is normal in HCR.

VII. Insert H_1, H_1Q_1 and H_1R in the lattice.

9. Since $P_1 \leqq P$, and H is maximal in the ordering relative to $\mathfrak{H}(ccl(A))$, by lemma 6.9, H contains every subgroup of type I. Since $Z(P)$ is of type I, $Z(P) \leqq H$, so that $Z(P) \leqq H_1$. Form $Z(H_1)$, and observe that $Z(H_1)$ contains $Z(P)$.

10. Form the subgroup $W = \{w \mid w \in Z(H_1), w^p = 1\}$, and $V = W \cap Z(P)$.

VIII. Insert $Z(H_1)$, $Z(P)$, W, WQ_1 (Q_1 is normal in HCR), and V in the lattice. Insert 1 in the lattice.

This completes the construction of the lattice.

Diagram II has put in view the relevant groups under discussion. We now turn to the relevant homomorphisms. $Z(H_1)$ is characteristic in H_1, and W is characteristic in $Z(H_1)$, hence is characteristic in H_1. Thus, WQ_1/Q_1 is characteristic in H_1Q_1/Q_1. Since $H_1Q_1 \lhd HCR$, we have $H_1Q_1/Q_1 \lhd HCR/Q_1$, and hence $WQ_1/Q_1 \lhd HCR/Q_1$. In the group HCR, we have the following chain of subgroups, each normal in HCR: $1 \leqq Q_1 < Q_1W \leqq Q_1H_1 < HCR$. Since Q_1 and Q_1W are both normal in HCR, HCR induces automorphisms on Q_1W/Q_1 by conjugation. $Q_1W/Q_1 \cong W$, and W, from its definition, is an elementary abelian p-group. Thus we have a homomorphism from HCR into $\mathrm{Aut}(Q_1W/Q_1)$. Let L be the kernel of this homomorphism.

Lemma 7.1. $L = H_1Q_1$.

Proof. Since $W \leqq Z(H_1)$, H_1 is contained in L. Q_1 is obviously contained in L, so we have $L \geqq H_1Q_1$. Let gQ_1 be any q-element of HCR/Q_1 which centralizes WQ_1/Q_1. Then gQ_1 centralizes VQ_1/Q_1, that is, if $v \in V$, then $g^{-1}vg = vg_1$, where g_1 lies in Q_1. But V is contained in H, and H and Q_1 commute elementwise, being normal subgroups of HCR of coprime orders. Now vg_1, being a conjugate of v, must be a p-element. Since v and g_1 commute, g_1 must be a p-element. Since $g_1 \in Q_1$, we have $g_1 = 1$, that is, g centralizes V.

V is characteristic in $Z(P)$, hence of type I, and in particular, is $\mathfrak{A}$-invariant. By lemma 5.3, $N_G(V)$ has a normal p-complement, say E. Let $\tilde{Q}$ be a q-Sylow subgroup of L. By the previous paragraph $\tilde{Q} \leqq N_G(V)$, since it is even contained in $C_G(V)$, so $\tilde{Q} \leqq E$. Also, $H \leqq N_G(V)$, since V is normal in P and H is contained in P. Finally, $\tilde{Q}$ normalizes H, since H is normal in HCR. Hence $[\tilde{Q}, H] \leqq H$. But $[\tilde{Q}, H] \leqq [E, H] \leqq E$, since $E \lhd N(V)$. Hence, $[\tilde{Q}, H] \leqq H \cap E = 1$, since H and E have coprime order; that is, $\tilde{Q}$ centralizes H. But then by condition 5 in the definition of the subgroup K of type III with respect to N, we find that $\tilde{Q} = Q_1$.

We have shown that Q_1 is a q-Sylow subgroup of L. Hence, L/Q_1 is a normal p-subgroup of HCR/Q_1. Since H_1Q_1/Q_1 is the maximal normal p-subgroup of HCR/Q_1, it follows that $L = H_1Q_1$, as asserted.

LEMMA 7.2. *HC/H_1 is represented faithfully on W and an element cH_1 of order $r = p^m$ has the minimal equation $(x-1)^r = 0$ on W.*

PROOF. We first verify that H_1 is normal in HC, in order that HC/H_1 make sense. Since $H_1Q_1 \lhd HCR$ by lemma 7.1, we have $H_1Q_1 \cap HC \lhd HC$. Clearly, $H_1Q_1 \cap HC \geqq H_1$, since $HC \geqq H_1$. Also, however, $H_1Q_1 \cap HC \leqq H_1$, since Q_1 is a q-group. Thus $H_1 \lhd HC$.

In order to prove this lemma, we make use of the following diagram:

$$\begin{array}{ccccc} HCQ_1/H_1Q_1 & \xrightarrow{\varrho} & \mathrm{Aut}(WQ_1/Q_1) & & WQ_1/Q_1 \\ \lambda\downarrow & & \downarrow\mu^* & & \downarrow\mu \\ HC/H_1 & \xrightarrow[\varrho^*]{} & \mathrm{Aut}(W) & & W \end{array}$$

ϱ is the homomorphism which maps cH_1Q_1 onto the automorphism $\varrho(c)$ defined by: $\varrho(c)(wQ_1) = c^{-1}wcQ_1$. λ is the isomorphism: $\lambda(cH_1Q_1) = cH_1$. μ is the isomorphism: $\mu(wQ_1) = w$. μ^* is the isomorphism of $\mathrm{Aut}(WQ_1/Q_1)$ and $\mathrm{Aut}(W)$ induced by μ: if $\sigma \in \mathrm{Aut}(WQ_1/Q_1)$, then $\mu^*(\sigma)(w) = \mu(\sigma(wQ_1))$. Finally, we define $(\varrho^*(cH_1))(w) = c^{-1}wc$. One verifies directly that the diagram is commutative. Since ϱ^* obviously extends to a map of the group ring of HC/H_1 over integers mod p into the endomorphism ring of W and ϱ extends similarly, the question of the minimal polynomials can be thrown onto the behavior of ϱ, which is what we do.

Lemma 7.1 states that HCR/H_1Q_1 is represented faithfully on WQ_1/Q_1. From the definition of H_1, HCR/H_1Q_1 has no non-trivial normal p-subgroups. A q-Sylow subgroup of HCR/H_1Q_1 is elementary abelian. Hence, HCR/H_1Q_1 is a p-soluble group with no non-trivial normal p-groups which is represented faithfully as linear transformations of the vector space WQ_1/Q_1. Theorem B of HALL and HIGMAN obtains, and the lemma is proved.

LEMMA 7.3. *$Z(HC/H_1)$ is cyclic.*

PROOF. Since λ: $HCQ_1/H_1Q_1 \cong HC/H_1$, it suffices to prove that $Z(HCQ_1/H_1Q_1)$ is cyclic. Now HCQ_1/H_1Q_1 induces automorphisms of RH_1/H_1Q_1. By the choice of R, RH_1/H_1Q_1 $(\cong RH/Q_1H)$ is a minimal

HCR/H_1Q_1-invariant subgroup of RH_1/H_1Q_1, and since RH_1/H_1Q_1 ($\cong R/Q_1$) is abelian by condition 4 in the definition of a subgroup of type III with respect to H, it follows that RH_1/H_1Q_1 is a minimal HCQ_1/H_1Q_1-invariant subgroup, that is, HCQ_1/H_1Q_1 operates irreducibly on RH_1/H_1Q_1.

Since H_1Q_1/Q_1 is the maximal normal p-group of HCR/Q_1, the group HCR/H_1Q_1 has no non-trivial normal p-groups. Hence, the group RH_1/H_1Q_1, which is the maximal normal q-group of HCR/H_1Q_1 (it is a q-Sylow subgroup of HCR/H_1Q_1) contains its centralizer in HCR/H_1Q_1 by 2.3.h. It follows directly from this fact that HCQ_1/H_1Q_1 operates faithfully on RH_1/H_1Q_1.

We have shown that HCQ_1/H_1Q_1 operates faithfully and irreducibly on RH_1/H_1Q_1, this last group being an elementary abelian q-group, so that it can be considered as a vector space over integers mod q. Lemma 3.11 applies, and the lemma is proved.

VIII. The Concluding Lemmas

The situation studied in sections 6 and 7 led to a lattice configuration together with certain homomorphisms. The study was carried on in such a manner that the choices involved in constructing the situation were all given explicitly. Since all the necessary preliminary results have been obtained, it is worthwhile to summarize the places in the argument where a choice has been made, before proving the concluding lemmas.

Consider the 6-tuple (A, H, K, Q, C, R). The only choices in the argument of section 7 occurred in the selection of these six subgroups, as follows:

1. Select a subgroup A of type I.
2. Select $H \in \mathfrak{H}(ccl(A))$, H maximal in the ordering $\leqq$ relative to $\mathfrak{H}(ccl(A))$, such that $P \cap N_G(H) = P_1$ is a p-Sylow subgroup of $N_G(H)$.
3. Select K a subgroup of type III with respect to H, and containing P_1.

(D) 4. Select a q-Sylow subgroup Q of K.

5. Select C, a conjugate of A contained in P_1.
6. Select R such that $Q_1 < R \leqq Q$ and such that RH/HQ_1 is a minimal normal subgroup of HCQ/HQ_1. (Here Q_1 is uniquely determined as the maximal normal q-subgroup of K.)

When these choices are made the subgroup H_1 is uniquely determined: $H_1 = HC \cap L$, where L is the maximal normal subgroup of HCR such that L/Q_1 is a p-group.

The next two lemmas take us to the heart of the matter.

LEMMA 8.1. *Suppose for some choice of A, H, K, and Q satisfying $D1-4$, we can find C $(=A^y)$ and R satisfying $D5$ and $D6$ such that $H_1 = HC \cap L < HC$. If $B \neq 1$ is any $\mathfrak{A}$-invariant normal subgroup of P contained in A, then $B^y = B_1$ is not contained in H_1.*

PROOF. Suppose false and that $B_1 \leqq H_1$ for suitable choice of $B \neq 1$. Then the p-subgroup CW normalizes B_1 and is thus contained in some p-Sylow subgroup S of $N_G(B_1)$. Since $B \lhd P$, $B_1 \lhd P^y$, so S is a p-Sylow subgroup of G,

too. By lemmas 5.2 and 5.3, $N_G(B_1)$ has a normal p-complement. Now $C \lhd P^y$ and $Z(C) \lhd P^y$ so by lemma 3.2, C and $Z(C)$ are normal in S, and hence are both normal in $T = \{S, P^y\}$. Hence $C/Z(C) \lhd T/Z(C)$ and $C/Z(C) \leqq Z(P^y/Z(C))$ since $[P^y, C] \leqq Z(C)$, C being conjugate to a subgroup of type I. By lemma 3.2, $T/Z(C)$ is p-normal, so by lemma 3.9, $C/Z(C) \leqq Z(P^y/Z(C))$ implies that $C/Z(C) \leqq Z(S/Z(C))$, or $[S, C] \leqq Z(C)$. Hence, $[S, C, C] = 1$, and in particular $[W, C, C] = 1$, a contradiction to lemma 7.2 since $H_1 < HC$.

LEMMA 8.2. *For any choice of A, H, K, and Q satisfying D1–4 above, there exist C and R satisfying D5 and D6 such that $H_1 = HC \cap L < HC$.*

PROOF. Lemma 8.2 will be proved by showing that if it were false, then the triple $(G, P, \mathfrak{A})$ would not have property $\mathfrak{N}$. To be very specific, we shall derive a contradiction from the following assumptions:

(E)
1. $(G, P, \mathfrak{A})$ has property $\mathfrak{N}$.
2. Theorem A is false for the triple $(G, P, \mathfrak{A})$.
3. G is of minimal order with properties 1 and 2.
4. Lemma 8.2 is false.

To prove lemma 8.2, we first prove three other lemmas.

LEMMA 8.3. *Under the assumptions (E), $V(ccl(A), H) = V(ccl(A), P_1)$.*

PROOF. Lemma 8.3 demands that we show that if C is a conjugate of A contained in P_1, then C is even contained in H.

For convenience, put $V(ccl(A), H) = V(H)$ and $V(ccl(A), P_1) = V(P_1)$. Let $C_1, \ldots, C_s$ be the distinct conjugates of A contained in P_1. Since lemma 8.2 is assumed false, $HC_i = H_1$ for each i, and any choice of R. H_1 depends on C_i, of course, but we are assuming that it does not depend on R, and that $HC_i = H_1$ for all choices of R. Hence Q/Q_1 normalizes HC_iQ_1/Q_1, since Q/Q_1 is generated by the R/Q_1. Since this is true for all i, we see that Q/Q_1 normalizes JQ_1/Q_1, where $J = \{H, C_1, \ldots, C_s\}$. Lemma 8.3 is the statement that $H = J$.

J is normal in P_1, since P_1 normalizes H and permutes the C's among themselves. Suppose by way of contradiction that $H < J$. By lemma 6.1, a q-Sylow subgroup of $N_G(J)$ centralizes J. Since JQ_1 is normalized by Q by the preceding paragraph, and since $Q_1 \lhd K$, we find that $JQ_1 \lhd \{P_1, Q\} = K$, so that $JQ_1/Q_1 \lhd K/Q_1$. This contradicts condition 6 in the definition of a subgroup of type III with respect to H, and completes the proof of lemma 8.3.

LEMMA 8.4. *Under the assumptions (E), $P_1 = P$.*

PROOF. Suppose $P_1 < P$. Then by 2.3.e, P_1 is contained as a proper normal subgroup of some p-subgroup P_2. P_2 normalizes $V(P_1)$, since it permutes the C's among themselves. Since $V(P_1) = V(H)$ by lemma 8.3, $V(P_1)$ is normal in K, since the elements of K induce automorphisms of H which permute the C's among themselves. By lemma 5.9, with A playing the role of A, D playing the role of H, and R playing the role of Q, Q does not centralize A. A fortiori, Q does not centralize $V(P_1)$. But then $V(P_1) \in \mathfrak{H}(ccl(A))$ and is larger than H in the ordering $\ll$ relative to $\mathfrak{H}(ccl(A))$, since $\alpha_p(P_2) > \alpha_p(P_1) = \alpha_p(N(H))$.

This completes the proof of lemma 8.4.

LEMMA 8.5. *Under the assumptions* (E), $V(H)$ *is an* $\mathfrak{N}$*-invariant normal subgroup of* P *such that* $A_G(V(H))$ *is not a* p*-group, so that* $(G, P, \mathfrak{N})$ *does not have property* $\mathfrak{N}$.

PROOF. Since $V(H) = V(P_1)$ and $P_1 = P$ by lemmas 8.3 and 8.4, $V(P) = V(H)$ is normal in P. By lemma 5.8, $V(P)$ is $\mathfrak{N}$-invariant. Q does not centralize $V(H)$, since it does not centralize A. But Q does normalize $V(H)$. This completes the proof of lemma 8.5, and proves lemma 8.2 as well.

The object of lemmas 8.1 through 8.5 has been to obtain information about the group HC independent of its embedding in G. From this point to the end of the proof of Theorem A, the relevant portion of diagram II is the heavily-inked sublattice involving only p-subgroups.

LEMMA 8.6. *If* A *is an abelian subgroup of type I which is not elementary abelian, then* $H_1 = HC$, *for any choice of* $H, K, Q, C,$ *and* R, *which satisfy* $D1-6$.

PROOF. If $C \leqq H$, then $H_1 = H = HC$, so we are done. Suppose by way of contradiction that $C \not\leqq H$, and that $H_1 < CH$. By lemma 7.3, $Z(HC/H_1)$ is cyclic. Since $HC/H_1 \cong C/C \cap H_1$, and C is a conjugate of the abelian p-group A, we conclude that HC/H_1 is cyclic of order p^d, $d \geqq 0$, so the lemma demands that we show $d = 0$.

CASE 1. $d = 1$. Since A is not elementary abelian, neither is C, so $\varphi(C) \neq 1$. Since $C \cap H_1$ is of index $p^d = p$ in C, we conclude that $\varphi(C)$ is contained in H_1. This is not possible by lemma 8.1.

CASE 2. $d > 1$. We choose $c \in C$ so that cH_1 is a cyclic generator for CH/H_1. We then choose a subgroup W_1 of W, and a basis $w_1, \ldots, w_t$ of W_1, where $t = p^d$, with respect to which c induces an automorphism with matrix in Jordan canonical form

$$I + N = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 1 & 1 \\ 0 & \ldots & 0 & 1 \end{pmatrix}.$$

Lemma 7.2 guarantees that we can find such a W_1.

The matrix of c^p is $(I+N)^p = I + N^p$, so that c^p centralizes $w_{t-(p-1)}$, $w_{t-(p-2)}, \ldots, w_t$, since N^p has zeroes in its last p rows. Since $p \geqq 3$, c^p centralizes the subgroup $W_2 = \{w_{t-2}, w_{t-1}, w_t\}$. But c^p and $C \cap H_1$ generate a subgroup which contains the p-th power of every element in C, so we conclude that $\varphi(C)$ centralizes W_2.

The group CW_2 is a p-subgroup contained in HC, and it normalizes $\varphi(C)$, since it even centralizes it. Enlarge CW_2 to a p-Sylow subgroup S_1 of $N_G(\varphi(C))$. By lemmas 5.2 and 5.3, $N_G(\varphi(C))$ has a normal p-complement. C is normal in the p-Sylow subgroup P^y of G, and since $\varphi(C)$ is characteristic in C, P^y is a p-Sylow subgroup of $N_G(\varphi(C))$. By lemma 3.2, C is normal in S_1, so that

$[S_1, C] \leqq C$. Since C is abelian $[S_1, C, C] = 1$. Since W_2 is a subgroup of S_1, we also have $[W_2, C, C] = 1$. But $[w_{t-2}, c, c] = w_t \neq 1$, and this contradiction shows that case two is not possible.

Since $d = 1$ and $d > 1$ are both impossible, we conclude that $d = 0$, that is, $H_1 = HC$.

LEMMA 8.7. *If A is any abelian subgroup of type I, then A is elementary abelian.*

PROOF. This is an immediate consequence of lemma 8.6, in conjunction with lemma 8.2.

LEMMA 8.8. *If A is a maximal abelian normal subgroup of P which is $\mathfrak{A}$-invariant, then $H_1 = HC$, for any choice of H, K, Q, C, and R which satisfy $D1$–6.*

PROOF. If $C \leqq H$, then $H_1 = HC = H$, and we are done. Suppose by way of contradiction that $H_1 < HC$. Since $CH/H_1 \cong C/C \cap H_1$, and since C is abelian, we conclude that CH/H_1 is abelian. By lemma 8.3, CH/H_1 is cyclic. By lemma 8.7, A is elementary abelian, so $C = A^y$ is also elementary abelian. Hence, HC/H_1 is of order 1 or p, so must be of order p, since $H_1 < HC$. Since $A \lhd P$, we have $C \lhd P^y$. Since A is a maximal abelian normal subgroup of P, $Z(P) \leqq A$. Hence $Z(P^y) \leqq C$.

It follows from lemma 8.1 that $Z(P^y) \not\leqq H_1$, so we choose an element $c \in Z(P^y)$ such that cH_1 is a cyclic generator for HC/H_1. In W, choose a subgroup W_1, and a basis $w_1, \ldots, w_p$ for W_1 such that c induces an automorphism on W_1 with matrix in Jordan canonical form

$$\begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & & 1 & 1 \\ 0 & \ldots & & 0 & 1 \end{pmatrix},$$

which is possible by lemma 7.2. Since c centralizes w_p, we conclude that w_p centralizes all of C. Since C is a maximal abelian normal subgroup which is conjugate to the $\mathfrak{A}$-invariant normal subgroup A of P, we conclude that w_p lies in C, by lemma 5.7. Put $D = \{c, w_p\}$. We see that w_{p-1} normalizes D but does not centralize c. Since w_{p-1} centralizes $C \cap H_1$, it follows that w_{p-1} normalizes $\{C \cap H_1, D\}$, which is equal to C, since $C \cap H_1$ and c generate C and w_p is contained in C. Hence, the group $E = \{C, w_{p-1}\}$ is a p-group contained in $N_G(C)$.

Enlarge the group E to a p-Sylow subgroup of $N_G(C)$, say S. Then S is a p-Sylow subgroup of G, too, since $C \lhd P^y$. $N_G(C)$ has a normal p-complement by lemmas 5.2 and 5.3. By lemma 3.2, we conclude that $C \lhd S$.

Since $N_G(C)$ has a normal p-complement, it is p-normal by lemma 3.2. Since $Z(P^y) < S$, and P^y is a p-Sylow subgroup of $N_G(C)$, $Z(P^y) = Z(S)$, by p-normality. But w_{p-1} is an element of S which does not centralize the element c of $Z(P^y)$. This contradiction completes the proof of lemma 8.8.

LEMMA 8.9. *If A is any maximal abelian normal subgroup of P, then A is not $\mathfrak{A}$-invariant.*

PROOF. This is an immediate consequence of lemma 8.8, in conjunction with lemma 8.2.

Since none of the maximal abelian normal subgroups of P is $\mathfrak{A}$-invariant, lemma 3.7 guarantees that P contains subgroups of type II. Since subgroups of type II are also subgroups of type I, the results of sections 6 and 7 are valid.

LEMMA 8.10. *If A is any subgroup of type II, then either $H_1 = HC$ or HC/H_1 is non-abelian for any choice of H, K, Q, C, and R satisfying $D1$—6.*

PROOF. This is a consequence of lemma 8.1, since if $H_1 < HC$ and HC/H_1 is abelian, then $C' \leqq H_1$. Since $C' \neq 1$, this contradicts lemma 8.1.

LEMMA 8.11. *If A is any subgroup of type II, then $H_1 = HC$ for any choice of H, K, Q, C $(= A^y)$ and R satisfying $D1$—6.*

PROOF. Assume by way of contradiction that $H_1 < HC$. Then lemma 8.10 says that HC/H_1 is not abelian. By lemma 8.1, $Z(P^y) \leqq H_1$, and $Z(P^y)$ is elementary abelian by lemma 8.7. Now $Z(C)/Z(C) \cap H_1 \cong Z(C)H_1/H_1 \leqq Z(CH_1/H_1)$, and since $Z(CH/H_1)$ is cyclic by lemma 7.3 and $Z(C)$ is elementary abelian by lemma 8.7, we conclude that $Z(C)/Z(C) \cap H_1$ has order 1 or p. But $Z(P^y) \leqq Z(C)$ by condition 5 in the definition of a subgroup of type II, and $Z(P^y) \leqq H_1$, so we must have $Z(C)/Z(C) \cap H_1$ of order p, and

$$(***) \qquad Z(C) = \{c, Z(C) \cap H_1\}$$

for some $c \in Z(P^y)$.

Since HC/H_1 is not abelian, there exist elements $a, b \in C - H_1$ with $[a, b] \in C' \leqq Z(C)$, but $[a, b] \in' H_1$. Then $[a, b] = c^k d$, with $0 < k < p$, and $d \in Z(C) \cap H_1$, and by altering the choice of c if necessary, we assume that $k = 1$. Now we form the group $X = \{a, b, H_1\}/H_1 \leqq CH_1/H_1$. Then $[aH_1, bH_1] = cH_1$ lies in $Z(X)$. Also $(cH_1)^p = H_1$, and since $\varphi(C) \leqq Z(C)$ by condition 3 in the definition of a subgroup of type II, we have $(aH_1)^p = a^p H_1 \in Z(C)H_1/H_1 = \{c, H_1\}/H_1$, and similarly for b. Therefore $o(X) = p^3$.

X is represented faithfully on W since X is a subgroup of HC/H_1 which is represented faithfully on W by lemma 7.2. Let V_1 be a subgroup of W of minimal possible order which is X-invariant and on which X is represented faithfully.

The number of subgroups of index p in V_1 is $\equiv 1 \pmod{p}$ [2, p. 152]. X induces automorphisms of V_1 which permute these subgroups among themselves. Hence, X must leave at least one of them fixed, say V_2. Since $o(V_1) > o(V_2)$, X is not represented faithfully on V_2. It follows that V_2 centralizes c, since $\{c, H_1\}/H_1 = Z(X)$ is contained in every non-trivial normal subgroup of X. It follows that V_2 centralizes $\{c, Z(C) \cap H_1\} = Z(C)$, by (***). The group $\{V_2, C\}$ is contained in the p-group HC, so is a p-group. By lemmas 5.2 and 5.6, with $Z(C)$ in the role of B and C in the role of C and V_2 in the role of T, we conclude that $[V_2, C] \leqq Z(C)$. Put $W_2 = [V_2, C]$, and $W_1 = [V_1, c]$. Now V_1 and V_2 are both X-invariant, and $[V_1 : V_2] = p$. Hence X is represented as

linear transformations on V_1/V_2, which is the cyclic group of order p. Since X is a p-group, this representation is trivial. In group-theoretic terms, this says that $[v_1, x] \in V_2$ for all $v_1 \in V_1$ and $x \in X$. In linear transformation terms, this says that $v_1(I - x) \in V_2$ for all $v_1 \in V_1$ and $x \in X$. We apply lemma 3.5 to conclude that $W_1 \leqq W_2$.

Since $W_2 \leqq Z(C)$, we conclude that $W_1 \leqq Z(C)$. Now V_1 centralizes H_1 and hence $Z(C) \cap H_1$. Since $[V_1, c] = W_1 \leqq Z(C)$, it follows from (***) that V_1 normalizes $Z(C)$. Since X is represented faithfully on V_1, V_1 possesses an element v which does not centralize c.

The group $Z(C)V_1$ is contained in the p-group HC, so is a p-group. Enlarge $Z(C)V_1$ to a p-Sylow subgroup S of $N_G(Z(C))$. By lemmas 5.2 and 5.3, $N_G(Z(C))$ has a normal p-complement. Now $Z(P^y) \leqq Z(C)$, since C is conjugate to A, and A is of type II. Hence $Z(P^y) \leqq S$. Also P^y is a p-Sylow subgroup of $N_G(Z(C))$. Since $N_G(Z(C))$ has a normal p-complement, it is p-normal by lemma 3.2. Since $Z(P^y) \leqq S$, we have $Z(P^y) = Z(S)$. But $c \in Z(P^y)$ and v is an element of S which does not centralize c. This contradiction completes the proof of lemma 8.11 and at the same time completes the proof of Theorem A, since lemma 8.11 is in contradiction with lemma 8.2.

Bibliography

[1] BAER, R.: Classes of finite groups and their properties. Ill. J. Math. **1**, 115–187 (1957). — [2] BURNSIDE, W.: Theory of groups of finite order, 2nd ed. New York: Dover 1955. — [3] HALL, M.: The theory of groups. New York: MacMillan 1959. — [4] HALL, P.: A contribution to the theory of groups of prime-power order. Proc. Lon. Math. Soc. **36**, 29–95 (1933/34). — [5] HALL, P., and G. HIGMAN: On the p-length of p-soluble groups. Proc. Lon. Math. Soc. **6**, 1–42 (1956). — [6] THOMPSON, J.: Finite groups with fixed-point-free automorphisms of prime order. Proc. Nat. Acad. Sci. **45**, 578–581 (1959). — [7] WIELANDT, H.: p-Sylow-Gruppen und p-Faktor-Gruppen. J. reine angew. Math. **182**, 180–193 (1940). — [8] ZASSENHAUS, H.: The theory of groups, 2nd English ed. New York: Chelsea 1958.

Dept. of Math., University of Chicago, Chicago 37, Ill. (U.S.A.)

(Eingegangen am 22. Mai 1959)

Hodges, J. H.
Math. Zeitschr. 72, 355–361 (1960)

Some Determinantal Equations Over a Finite Field

By
John H. Hodges

1. Introduction and notation. Let $q=p^n$, p a prime, and let $GF(q)$ denote the finite field of order q. Let A be a nonsingular matrix of order m over $GF(q)$ and let t be an integer such that $1\leqq t\leqq m$. In this paper, we consider the problem of finding the number of solutions U, V of the determinantal equation

$$\begin{vmatrix} A & V \\ U & 0 \end{vmatrix} = \beta, \tag{1.1}$$

where β is an assigned number of $GF(q)$, and U and V are $t\times m$ and $m\times t$ matrices, respectively, over $GF(q)$.

For $t=1$, (1.1) reduces to an ordinary bilinear equation, for which the number of solutions is well known. In order to determine the number of solutions for arbitrary t, the determinantal equation (1.1) is reduced to a form which is shown to be related to a problem involving bilinear forms over $GF(q)$. This latter problem has been solved in [5]. By using the relationship between the two problems, the number of solutions of (1.1) is found (Theorem 1) for $\beta\neq 0$. Then, by using Theorem 1, the number of solutions of (1.1) is found (Theorem 2) for $\beta=0$. As a check on these theorems, it is shown in § 5 that the results which they give for $t=1$ and $t=m$ coincide with results which can be obtained directly. For $\beta=0$, the solutions of (1.1) may be restricted to those for which both U and V are of rank t. The number of such *proper* solutions is found in Theorem 3. In § 7 we consider the generalized determinantal equation

$$\sum_{i=1}^{k} \begin{vmatrix} A_i & V_i \\ U_i & 0 \end{vmatrix} = \beta, \tag{1.2}$$

where $k\geqq 1$, the A_i are nonsingular matrices of order m_i, U_i and V_i are $t_i\times m_i$ and $m_i\times t_i$ matrices, respectively, and $1\leqq t_i\leqq m_i$ for $i=1,\ldots,k$. Finally, in § 8, as an application of the results for the number of solutions of (1.2), we evaluate several generalized Gauss sums. Some of the methods of this paper are similar to those employed by L. Carlitz in considering the analogous problems for symmetric matrices in [2] and for skew matrices in [3].

In the present paper, numbers of $GF(q)$ will be denoted by lower case Greek letters $\alpha, \beta, \ldots$. Matrices with elements in $GF(q)$ will be denoted by italic capitals $A, B, \ldots$. $A(m, s)$ will denote a matrix of m rows and s columns.

The research for this paper was supported, in part, by National Science Foundation Research Grant G-2990.

We will need the following formula, which is a special case of a formula due to LANDSBERG [6], for the number $g(m, t, t)$ of $m \times t$ matrices ($t \leqq m$) of rank t over $GF(q)$:

$$g(m, t, t) = \prod_{i=0}^{t-1} (q^m - q^i). \tag{1.3}$$

In particular,

$$g_t = g(t, t, t) = \prod_{i=0}^{t-1} (q^t - q^i) \tag{1.4}$$

is the number of nonsingular matrices of order t over $GF(q)$.

2. THE SIMPLE DETERMINANTAL EQUATION. Let $\beta \in GF(q)$ be fixed and let $A = A(m, m) = (\alpha_{ij})$ with $|A| = \delta \neq 0$. If t is an arbitrary integer $\geqq 1$, let $N(m, t; \beta) = N(A; m, t; \beta)$ denote the number of pairs $U = U(t, m) = (\mu_{rj})$, $V = V(m, t) = (\nu_{is})$ of matrices over $GF(q)$ such that

$$\begin{vmatrix} A & V \\ U & 0 \end{vmatrix} = \beta. \tag{2.1}$$

There is no loss of generality in assuming that $t \leqq m$. Let $A^{-1} = (\beta_{ij})$, so that $\beta_{ij} = \delta^{-1} A_{ij}$, where (A_{ij}) is the adjoint of A. Then for each $i = 1, \ldots, m$ and $r = 1, \ldots, t$, define

$$\lambda_{ir} = \sum_{j=1}^{m} \beta_{ij} \mu_{rj}. \tag{2.2}$$

Consider the determinant in (2.1). For each $r = 1, \ldots, t$ subtract, from the $(m+r)$-th row of the determinant, λ_{ir} times the i-th row, for $i = 1, \ldots, m$. If the resulting determinant is denoted by $G = |\gamma_{ab}|$, then for any $a = m + r$ with $1 \leqq r \leqq t$ and any b such that $1 \leqq b \leqq m$, we have

$$\begin{cases} \gamma_{ab} = \mu_{rb} - \sum_{i=1}^{m} \lambda_{ir} \alpha_{ib} = \mu_{rb} - \sum_{i=1}^{m} \alpha_{ib} \left(\sum_{j=1}^{m} \beta_{ij} \mu_{rj} \right) \\ \qquad = \mu_{rb} - \sum_{j=1}^{m} \mu_{rj} \left(\sum_{i=1}^{m} \alpha_{ib} \beta_{ij} \right). \end{cases} \tag{2.3}$$

But $\beta_{ij} = \delta^{-1} A_{ij}$, so that the inner sum in the right member of (2.3) is equal to

$$\sum_{i=1}^{m} \alpha_{ib} \beta_{ij} = \delta^{-1} \sum_{i=1}^{m} \alpha_{ib} A_{ij} = \begin{cases} 0 & (j \neq b), \\ \delta^{-1} \delta = 1 & (j = b). \end{cases} \tag{2.4}$$

Thus, using (2.4) in (2.3), we get, for $m+1 \leqq a = m+r \leqq m+t$ and $1 \leqq b \leqq m$,

$$\gamma_{ab} = \mu_{rb} - \mu_{rb} = 0. \tag{2.5}$$

Furthermore, for $m+1 \leqq a = m+r \leqq m+t$ and $m+1 \leqq b \leqq m+t$ so that $b = m+s$ with $1 \leqq s \leqq t$, we have

$$\gamma_{ab} = -\sum_{i=1}^{m} \lambda_{ir} \nu_{is} = -\sum_{i=1}^{m} \left(\sum_{j=1}^{m} \beta_{ij} \mu_{rj} \right) \nu_{is} = -\sum_{i,j=1}^{m} \beta_{ij} \mu_{rj} \nu_{is}. \tag{2.6}$$

In view of (2.5) and (2.6), it is clear that equation (2.1) reduces to

$$\left|\sum_{i,j=1}^{m} \beta_{ij}\mu_{rj}\nu_{is}\right| = (-1)^t \delta^{-1}\beta \qquad (r, s = 1, \ldots, t), \tag{2.7}$$

so that every solution of (2.1) is a solution of (2.7) and conversely. Therefore, $N(m, t; \beta)$, the number of solutions of (2.1), is also equal to the number of solutions of (2.7).

3. A RELATED PROBLEM. Let $A = A(m, k)$ and $B = B(s, t)$ have ranks r and z, respectively, where $z \leqq r$, and let $N_t^s(A, B)$ denote the number of pairs $U = U(s, m)$, $V = V(k, t)$ of matrices over $GF(q)$ such that

$$UAV = B. \tag{3.1}$$

The value of $N_t^s(A, B)$ has been found in [5]. We shall only consider the case where A and B are square and nonsingular, that is, $m = k = r$ and $s = t = z$. Then by [5, Theorem 3], we have explicitly

$$N_t^t(A, B) = q^{t(2m-t-1)/2} \prod_{i=1}^{t} (q^{m-i+1} - 1) = q^{t(m-t)}\, g(m, t, t), \tag{3.2}$$

where $g(m, t, t)$ is given by (1.3). It is easy to show that the number of square matrices B of order t and determinant $\beta \neq 0$ is $g_t/(q-1)$, where g_t, the number of nonsingular matrices of order t, is given by (1.4). If we take $A = A(m, m) = (\alpha_{ij})$, $U = U(t, m) = (\mu_{ri})$ and $V = V(m, t) = (\nu_{js})$ in (3.1), then

$$UAV = \left(\sum_{i,j=1}^{m} \alpha_{ij}\mu_{ri}\nu_{js}\right), \qquad (r, s = 1, \ldots, t). \tag{3.3}$$

Let B range over the non-singular matrices of order t and determinant $\beta \neq 0$. Since by (3.2), the number $N_t^t(A, B)$, as a function of B, depends only on the order of B, it follows that the number of solutions of the equation

$$|UAV| = \left|\sum_{i,j=1}^{m} \alpha_{ij}\mu_{ri}\nu_{js}\right| = \beta \neq 0, \qquad (r, s = 1, \ldots, t), \tag{3.4}$$

is given by multiplying $N_t^t(A, B)$ by $g_t/(q-1)$.

4. SOLUTIONS OF THE SIMPLE DETERMINANTAL EQUATION. Returning to (2.1), or equivalently to (2.7), we take $\beta \neq 0$. Then comparing (2.7) with (3.4), with A replaced by $(A')^{-1}$ in the latter, we get

THEOREM 1. *If $A = A(m, m)$ is nonsingular and $\beta \neq 0$, then the number $N(m, t; \beta)$ of pairs $U(t, m)$, $V(m, t)$ of matrices over $GF(q)$ which satisfy (2.1), is given by*

$$N(m, t; \beta) = \frac{q^{t(m-t)}\, g_t\, g(m, t, t)}{q-1}, \tag{4.1}$$

where $g(m, t, t)$ and g_t are given by (1.3) and (1.4) respectively.

To get $N(m, t; 0)$ we use the relationship

$$N(m, t; 0) + \sum_{\beta \neq 0} N(m, t; \beta) = q^{2mt}, \tag{4.2}$$

which is obtained by letting U and V in (2.1) range through all matrices of the appropriate sizes and counting the number of times which each β appears as the value of the determinant. Since, by (4.1), $N(m, t; \beta)$ has the same value for all $\beta \neq 0$, substituting this value into (4.2) and solving for $N(m, t; 0)$ proves

THEOREM 2. *If $A = A(m, m)$ is nonsingular and $\beta = 0$, then the number $N(m, t; 0)$, of pairs $U(t, m)$, $V(m, t)$ of matrices over $GF(q)$ which satisfy* (2.1), *is given by*

$$N(m, t; 0) = q^{2mt} - q^{t(m-t)} g_t g(m, t, t), \tag{4.3}$$

where $g(m, t, t)$ and g_t are given by (1.3) *and* (1.4), *respectively.*

5. SOME SPECIAL CASES. As a partial check on (4.1) and (4.3) we consider some special cases of (2.1). First, taking $t = 1$, we see that (4.1) becomes, for $\beta \neq 0$,

$$N(m, 1; \beta) = q^{m-1}(q^m - 1), \tag{5.1}$$

and (4.3) becomes

$$N(m, 1; 0) = q^{m-1}(q^m + q - 1). \tag{5.2}$$

But, for $t = 1$, with $\mu_{1j} = \mu_j$ and $\nu_{i1} = \nu_i$, equation (2.7) reduces to

$$\sum_{i,j=1}^{m} \beta_{ij} \mu_j \nu_i = -\delta^{-1} \beta, \tag{5.3}$$

where $|\beta_{ij}| \delta \neq 0$. After reducing (β_{ij}) to the identity matrix of order m by means of non-singular linear transformations of the μ's and ν's, it can easily be proved, by induction on m in the resulting bilinear equation which is equivalent to (5.3), that (5.1) and (5.2) actually give the number of solutions of (5.3) for $\beta \neq 0$ and $\beta = 0$, respectively. For q odd, also see [4; p. 47].

Secondly, taking $t = m$, (4.1) becomes, for $\beta \neq 0$,

$$N(m, m; \beta) = (g_m)^2/(q - 1), \tag{5.4}$$

and (4.3) reduces to

$$N(m, m; 0) = q^{2m^2} - (g_m)^2. \tag{5.5}$$

When $t = m$, equation (2.1) becomes

$$|U| \, |V| = (-1)^m \beta. \tag{5.6}$$

If $\beta \neq 0$, (5.6) implies that U and V are both nonsingular. For each of the g_m nonsingular U's, there are $g_m/(q-1)$ V's such that $|V| = (-1)^m \beta / |U|$, so that the number of solutions of (5.6) is given correctly by (5.4). If $\beta = 0$, then (5.6) implies that $|U| = |V| = 0$ or $|U| = 0$ and $|V| \neq 0$ or $|U| \neq 0$ and $|V| = 0$. The numbers of pairs U, V satisfying these three mutually exclusive conditions are $(q^{m^2} - g_m)^2$, $(q^{m^2} - g_m) g_m$, and $g_m (q^{m^2} - g_m)$ respectively, and the sum of these numbers is equal to $N(m, m; 0)$ as given by (5.5).

6. PROPER SOLUTIONS OF THE SIMPLE DETERMINANTAL EQUATION. When $\beta = 0$, a solution U, V of (2.1) is called *proper* if rank U = rank $V = t$. If U and V are of the appropriate sizes and either rank $U < t$ or rank $V < t$, then the pair U, V is clearly an improper (not proper) solution of (2.1). Since

the numbers of pairs U, V such that rank $U<t$ and rank $V<t$, rank $U=t$ and rank $V<t$, rank $U<t$ and rank $V=t$ are $\{q^{mt}-g(m,t,t)\}^2$, $g(m,t,t)\{q^{mt}-g(m,t,t)\}$, $\{q^{mt}-g(m,t,t)\}\,g(m,t,t)$, respectively, the sum of these three quantities is the number of improper solutions of (2.1). The total number $N(m,t;0)$ of solutions of (2.1) for $\beta=0$ is given by (4.3), so that subtracting the number of improper solutions from $N(m,t;0)$ gives

THEOREM 3. *If $A=A(m,m)$ is nonsingular and $\beta=0$, then the number $N'(m,t;0)$, of proper solutions $U(t,m)$, $V(m,t)$ of the determinantal equation (2.1), is given by*

$$N'(m,t;0)=g(m,t,t)\{g(m,t,t)-q^{t(m-t)}g_t\}, \tag{6.1}$$

where $g(m,t,t)$ and g_t are given by (1.3) and (1.4) respectively.

7. THE GENERALIZED DETERMINANTAL EQUATION. Let $\beta\in GF(q)$ and $m_1,\ldots,m_k$ and $t_1,\ldots,t_k$ be sets of $k\geqq 1$ positive integers with $1\leqq t_i\leqq m_i$ for $i=1,\ldots,k$. Let $N(m_1,\ldots,m_k;t_1,\ldots,t_k;\beta)$ denote the number of solutions $U_1,\ldots,U_k, V_1,\ldots,V_k$ of the equation

$$\sum_{i=1}^{k}\begin{vmatrix}A_i & V_i\\ U_i & 0\end{vmatrix}=\beta, \tag{7.1}$$

where A_i is a nonsingular matrix of order m_i and $U_i=U_i(t_i,m_i)$, $V_i=V_i(m_i,t_i)$ for $i=1,\ldots,k$. For $k=1$, $N(m_1,t_1;\beta)$ is given by (4.1) and (4.3). For $k=2$, it is evident that

$$\left.\begin{aligned}N(m_1,m_2;t_1,t_2;\beta)&=\sum_{\beta_1+\beta_2=\beta}N(m_1,t_1;\beta_1)\,N(m_2,t_2;\beta_2)\\&=\sum_{\beta_1}N(m_1,t_1;\beta_1)\,N(m_2,t_2;\beta-\beta_1)\\&=N(m_1,t_1;\beta)\,N(m_2,t_2;0)+\\&\quad+\sum_{\beta_1\neq\beta}N(m_1,t_1;\beta_1)\,N(m_2,t_2;\beta-\beta_1).\end{aligned}\right\} \tag{7.2}$$

Since $\sum_{\beta_1}N(m_1,t_1;\beta_1)=q^{2m_1t_1}$ and, in view of (4.1), $N(m_2,t_2;\beta-\beta_1)=N(m_2,t_2;1)$ for $\beta\neq\beta_1$, (7.2) becomes

$$\left.\begin{aligned}N(m_1,m_2;t_1,t_2;\beta)&=N(m_1,t_1;\beta)\,N(m_2,t_2;0)+\\&\quad+N(m_2,t_2;1)\{q^{2m_1t_1}-N(m_1,t_1;\beta)\}.\end{aligned}\right\} \tag{7.3}$$

Take $\beta\neq 0$, so that $N(m_1,t_1;\beta)=N(m_1,t_1;1)$. Then, the coefficient of $N(m_1,t_1;1)$ on the right side of (7.3) is

$$N(m_2,t_2;0)-N(m_2,t_2;1)=q^{2m_2t_2}-qN(m_2,t_2;1),$$

which follows from (4.1) and (4.2). Substituting this into (7.3) and simplifying gives

THEOREM 4. *If $\beta\neq 0$, the number of solutions U_1, U_2, V_1, V_2 of (7.1) with $k=2$ is given by*

$$\left.\begin{aligned}N(m_1,m_2;t_1,t_2;\beta)&=q^{2m_2t_2}N(m_1,t_1;1)+q^{2m_1t_1}N(m_2,t_2;1)-\\&\quad-qN(m_1,t_1;1)\,N(m_2,t_2;1),\end{aligned}\right\} \tag{7.4}$$

where $N(m,t;1)$ is given by (4.1).

Using Theorem 4, we can prove, by induction on k, the general

THEOREM 5. *If $\beta \neq 0$, the number $N_k = N(m_1, \ldots, m_k; t_1, \ldots, t_k; \beta)$ of solutions $U_1, \ldots, U_k$, $V_1, \ldots, V_k$ of (7.1) for $k \geqq 1$ is given by*

$$(7.5)\quad N_k = q^{2s_k} \sum_{i=1}^{k} (-q)^{i-1} \sum_{j_1, \ldots, j_i} q^{-2(m_{j_1} t_{j_1} + \cdots + m_{j_i} t_{j_i})} N(m_{j_1}, t_{j_1}; 1) \ldots N(m_{j_i}, t_{j_i}; 1),$$

where $s_k = \sum_{i=1}^{k} m_i t_i$, the inner sum is over all combinations of the k integers $1, 2, \ldots, k$ taken i at a time $(1 \leqq i \leqq k)$, and $N(m, t; 1)$ is given by (4.1).

We will not take the space to give the proof of this theorem here.

By analogy with the statement and proof of Theorem 2 we have

THEOREM 6. *If $\beta = 0$, the number $N(m_1, \ldots, m_k; t_1, \ldots, t_k; 0)$ of solutions $U_1, \ldots, U_k$, $V_1, \ldots, V_k$ of (7.1) for $k \geqq 1$ is given by*

$$(7.6)\quad N(m_1, \ldots, m_k; t_1, \ldots, t_k; 0) = q^{2s_k} - (q-1) N_k,$$

where $s_k = \sum_{i=1}^{k} m_i t_i$ and N_k is given by (7.5).

8. GENERALIZED GAUSS SUMS. Throughout this section we will assume that $q = p^n$ is odd. For $\alpha \in GF(q)$, put

$$(8.1)\quad e(\alpha) = e^{2\pi i t(\alpha)/p}, \qquad t(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{n-1}}.$$

Then it is easy to prove that

$$(8.2)\quad \sum_{\beta \in GF(q)} e(\alpha\beta) = \begin{cases} q & (\alpha = 0), \\ 0 & (\alpha \neq 0). \end{cases}$$

We recall the definition [*1*, §3] of the Gauss sum

$$(8.3)\quad G(\alpha) = \sum_{\beta} e(\alpha\beta^2),$$

for $\alpha \in GF(q)$, and the properties

$$(8.4)\quad \begin{aligned} G(\alpha) = \sum_{\beta} \psi(\beta)\, e(\alpha\beta) &= \psi(\alpha)\, G(1) \qquad (\alpha \neq 0); \\ G^2(1) &= \psi(-1)\, q, \end{aligned}$$

where $\psi(\alpha) = 0, 1, -1$ according to whether α is zero, a nonzero square, or a non-square of $GF(q)$. If $\alpha \in GF(q)$ and $m_1, \ldots, m_k$ and $t_1, \ldots, t_k$ are sets of positive integers with $1 \leqq t_i \leqq m_i$ for $i = 1, \ldots, k$, define

$$(8.5)\quad \begin{aligned} G_k(\alpha) &= G(m_1, \ldots, m_k; t_1, \ldots, t_k; \alpha) \\ &= \sum_{\beta} N(m_1, \ldots, m_k; t_1, \ldots, t_k; \beta)\, e(\alpha\beta^2). \end{aligned}$$

For $\alpha = 0$, it is obvious that $G_k(0) = q^{2s_k}$, where $s_k = \sum_{i=1}^{k} m_i t_i$. For $\alpha \neq 0$,

$$(8.6)\quad \begin{aligned} G_k(\alpha) = {} & N(m_1, \ldots, m_k; t_1, \ldots, t_k; 0) + \\ & + N(m_1, \ldots, m_k; t_1, \ldots, t_k; 1) \sum_{\beta \neq 0} e(\alpha\beta^2). \end{aligned}$$

If we substitute for $N(m_1, \ldots, m_k; t_1, \ldots, t_k; 0)$ in (8.6) its value in terms of $N_k = N(m_1, \ldots, m_k; t_1, \ldots, t_k; 1)$ as given by (7.6) and observe that, in view of (8.3), the summation on the right of (8.6) is equal to $G(\alpha) - 1 = \psi(\alpha) G(1) - 1$, we may state

THEOREM 7. *If $\alpha \in GF(q)$, q odd, and $G_k(\alpha)$ is the generalized Gauss sum defined by* (8.5), *then*

$$G_k(\alpha) = q^{2s_k} + \psi^2(\alpha)\{\psi(\alpha) G(1) - q\} N_k, \tag{8.7}$$

where $s_k = \sum_{i=1}^{k} m_i t_i$ and N_k is given by (7.5).

In similar fashion, if we define

$$\left\{ \begin{aligned} H_k(\alpha) &= H(m_1, \ldots, m_k; t_1, \ldots, t_k; \alpha) \\ &= \sum_{\beta} N(m_1, \ldots, m_k; t_1, \ldots, t_k; \beta)\, e(\alpha\beta), \end{aligned} \right. \tag{8.8}$$

where the m_i and t_i are as defined previously, then proceeding as in the proof of Theorem 7, using (8.2) instead of (8.3), we can prove

THEOREM 8. *If $\alpha \in GF(q)$, q odd, and $H_k(\alpha)$ is the generalized Gauss sum defined by* (8.8), *then*

$$H_k(\alpha) = q^{2s_k} - \psi^2(\alpha)\, q\, N_k, \tag{8.9}$$

where $s_k = \sum_{i=1}^{k} m_i t_i$ and N_k is given by (7.5).

Finally, defining

$$\left\{ \begin{aligned} H'_k(\alpha) &= H'(m_1, \ldots, m_k; t_1, \ldots, t_k; \alpha) \\ &= \sum_{\beta} \psi(\beta)\, N(m_1, \ldots, m_k; t_1, \ldots, t_k; \beta)\, e(\alpha\beta), \end{aligned} \right. \tag{8.10}$$

with the m_i and t_i as before, it is easy to prove

THEOREM 9. *If $\alpha \in GF(q)$, q odd, and $H'_k(\alpha)$ is the generalized Gauss sum defined by* (8.10) *,then*

$$H'_k(\alpha) = \psi(\alpha) N_k G(1) = G_k(\alpha) - H_k(\alpha), \tag{8.11}$$

where N_k is given by (7.5).

References

[1] CARLITZ, L.: The singular series for sums of squares of polynomials. Duke Math. J. **14**, 1105–1120 (1947). — [2] CARLITZ, L.: A problem involving quadratic forms in a finite field. Mathematische Nachrichten **11**, 135–142 (1954). — [3] CARLITZ, L.: Representations by skew forms in a finite field. Archiv der Math. **5**, 19–31 (1954). — [4] DICKSON, L. E.: Linear groups. Leipzig 1901. — [5] HODGES, JOHN H.: Representations by bilinear forms in a finite field. Duke Math. J. **22**, 497–510 (1955). — [6] LANDSBERG, GEORG: Über eine Anzahlbestimmung und eine damit zusammenhängende Reihe. J. reine angew. Math. **111**, 87–88 (1893).

Cornell Aeronautical Laboratory, Inc. of Cornell University, Buffalo 21, N.Y.(U.S.A.)

(Eingegangen am 10. März 1959)

Bagemihl, F.
Math. Zeitschr. 72, 362–366 (1960)

Planar sections of spatial sets

By
Frederick Bagemihl

We denote by $\mathscr{E}$ the linear continuum, and by $\mathscr{S}$ the set of points of three-dimensional Euclidean space provided with a rectangular Cartesian coordinate system whose axes are the x-, y-, and z-axes. The planes that are perpendicular to the x- (y-, z-)axis will be called x-(y-, z-)planes. If the union of a set of x-(y-, z-)planes intersects the x-(y-, z-)axis in a linear set with a certain property, we say, for brevity, that the set of planes possesses this property; for example, a set of measure zero of z-planes is a set of z-planes whose union intersects the z-axis in a linear set of measure zero.

A theorem proved by Sierpiński [*4*, p. 12, Theorem 9] may be stated as follows:

If $\mathscr{S}=E_1\cup E_2\cup E_3$, where every x-plane intersects E_1 in a finite number of points and every y-plane intersects E_2 in a finite number of points, then there exists a z-plane that intersects E_3 in $2^{\aleph_0}$ points.

We shall first obtain two generalizations of this result.

If A is a subset of a plane P, we say that A is a planar set that is everywhere of power $2^{\aleph_0}$ (everywhere of second category, everywhere of positive exterior measure) if, for every open rectangle R in the plane P, the set $A\cap R$ is of power $2^{\aleph_0}$ (second category, positive exterior measure).

Theorem 1. *If $\mathscr{S}=E_1\cup E_2\cup E_3$, where every x-plane intersects E_1 in a finite number of points and every y-plane intersects E_2 in less than $2^{\aleph_0}$ points, then every z-plane, with at most $\aleph_0$ exceptions, intersects E_3 in a planar set that is everywhere of power $2^{\aleph_0}$, everywhere of second category, and everywhere of positive exterior measure.*

Proof. Let r_1, r_2, s_1, s_2 be rational numbers with $r_1<r_2$ and $s_1<s_2$. For every $z\in\mathscr{E}$, let

$$S_z=\{(x, y, z)\colon r_1<x<r_2,\ s_1<y<s_2,\ x\in\mathscr{E},\ y\in\mathscr{E}\}\cap E_3.$$

Suppose that there exists an enumerable set Z of real numbers such that, for every $z\in Z$, $|S_z|<2^{\aleph_0}$. Then $\left|\bigcup_{z\in Z} S_z\right|<2^{\aleph_0}$, and hence there exists a real number y_0 between s_1 and s_2, and a set X of $2^{\aleph_0}$ real numbers between r_1 and r_2, such that, for every $x_0\in X$, none of the $\aleph_0$ points (x_0, y_0, z) with $z\in Z$ belongs to E_3. Moreover, these $\aleph_0$ points all lie in the plane $x=x_0$ which, by hypothesis, contains only a finite number of points of E_1, and therefore $\aleph_0$ of these points belong to E_2. But then the plane $y=y_0$ contains $2^{\aleph_0}\cdot\aleph_0=2^{\aleph_0}$ points of E_2, contrary to hypothesis.

Suppose that there exists an enumerable set Z of real numbers such that, for every $z\in Z$, S_z is a planar set of first category. Then the (orthogonal)

projection of $\bigcup_{z\in Z} S_z$ on the xy-plane is also a planar set of first category, and hence, according to a theorem of KURATOWSKI and ULAM [3], there exists a real number y_0 between s_1 and s_2, and a residual subset X of the interval (r_1, r_2) such that, for every $x_0 \in X$, none of the $\aleph_0$ points (x_0, y_0, z) with $z \in Z$ belongs to E_3. Since $|X| = 2^{\aleph_0}$, this leads, as before, to a contradiction of the hypothesis concerning E_2.

Suppose, finally, that there exists an enumerable set Z of real numbers such that, for every $z \in Z$, S_z is a planar set of measure zero. Appealing this time to a theorem of FUBINI [2], we again arrive at a contradiction of the hypothesis concerning E_2.

Hence, there exists a finite set Z_{r_1, r_2, s_1, s_2} of real numbers such that, for every $z \in \mathscr{E} - Z_{r_1, r_2, s_1, s_2}$ the projection of S_z on the xy-plane intersects the rectangle

$$\{(x, y, 0): r_1 < x < r_2,\ s_1 < y < s_2,\ x \in \mathscr{E},\ y \in \mathscr{E}\}$$

in a planar set of power $2^{\aleph_0}$, second category, and positive exterior measure. Let $Z^* = \bigcup Z_{r_1, r_2, s_1, s_2}$ where r_1, r_2, s_1, s_2 range over the set of rational numbers so that $r_1 < r_2$ and $s_1 < s_2$. Then $|Z^*| \leqq \aleph_0$, and for every $z_0 \in \mathscr{E} - Z^*$, the z-plane $z = z_0$ intersects E_3 in a planar set that is everywhere of power $2^{\aleph_0}$, everywhere of second category, and everywhere of positive exterior measure.

THEOREM 2. *If $\mathscr{S} = E_1 \cup E_2 \cup E_3$, where every x-plane intersects E_1 in a planar nowhere dense set of points and every y-plane intersects E_2 in less than $2^{\aleph_0}$ points, then every z-plane, with at most a set of first category of exceptions, intersects E_3 in a planar set that is everywhere of power $2^{\aleph_0}$, everywhere of second category, and everywhere of positive exterior measure.*

PROOF. Let r_1, r_2, s_1, s_2 be rational numbers with $r_1 < r_2$ and $s_1 < s_2$, and put

$$I = \{x: r_1 < x < r_2,\ x \in \mathscr{E}\}, \qquad J = \{y: s_1 < y < s_2,\ y \in \mathscr{E}\}.$$

For every $z \in \mathscr{E}$, define S_z as in the proof of Theorem 1.

Suppose that there exists an enumerable everywhere dense subset Z of some interval of $\mathscr{E}$ such that, for every $z \in Z$, $|S_z| < 2^{\aleph_0}$. If S is the projection of $\bigcup_{z\in Z} S_z$ on the xy-plane, then, as $|Z| = \aleph_0$, we have $|S| < 2^{\aleph_0}$. Hence, if Y is an enumerable everywhere dense subset of J, then, for every $y_0 \in Y$, the set

$$T_{y_0} = \{(x, y_0, 0): r_1 < x < r_2, x \in \mathscr{E}\} \cap S$$

is of power less than $2^{\aleph_0}$; and if T is the set of abscissas of the points constituting the projection of $\bigcup_{y_0\in Y} T_{y_0}$ on the x-axis, then $|T| < 2^{\aleph_0}$. Let $X = I - T$, and observe that $|X| = 2^{\aleph_0}$. Then, if $x \in X$, $y \in Y$, and $z \in Z$, the point $(x, y, z) \notin E_3$. If $y_0 \in Y$, the hypothesis concerning E_2 implies that if X_{y_0} is the set of those elements $x_0 \in X$ such that, for at least one $z \in Z$, the point (x_0, y_0, z) belongs to E_2, then $|X_{y_0}| < 2^{\aleph_0}$. Set $X^* = \bigcup_{y_0\in Y} X_{y_0}$. Then, since $|Y| = \aleph_0$, we have $|X^*| < 2^{\aleph_0}$, so that $X - X^*$ is not empty. Let $x_0 \in X - X^*$. Then the set

$$\{(x_0, y, z): y \in Y,\ z \in Z\}$$

is a subset of E_1 as well as of the plane $x=x_0$, but is not nowhere dense in this plane, contradicting the hypothesis concerning E_1.

Suppose that there exists an enumerable everywhere dense subset Z of some interval of $\mathscr{E}$ such that, for every $z\in Z$, S_z is a planar set of first category. If S is the projection of $\bigcup_{z\in Z} S_z$ on the xy-plane, then, as $|Z|=\aleph_0$, S is a planar set of first category. It follows from a theorem of KURATOWSKI and ULAM [3] that there exists an enumerable everywhere dense subset Y of J such that, for every $y_0\in Y$, the set

$$T_{y_0}=\{(x, y_0, 0): r_1<x<r_2,\ x\in\mathscr{E}\}\cap S$$

is a linear set of first category; and if T is the projection of $\bigcup_{y_0\in Y} T_{y_0}$ on the x-axis, then, as $|Y|=\aleph_0$, T is a linear subset of I of first category. Let $X=I-T$, and observe that $|X|=2^{\aleph_0}$. As before, we ultimately arrive at a contradiction of the hypothesis concerning E_1.

Suppose, finally, that there exists an enumerable everywhere dense subset Z of some interval of $\mathscr{E}$ such that, for every $z\in Z$, S_z is a planar set of measure zero. Appealing this time to a theorem of FUBINI [2], we again arrive at a contradiction of the hypothesis concerning E_1.

Hence, there exists a nowhere dense set Z_{r_1,r_2,s_1,s_2} of real numbers such that, for every $z\in\mathscr{E}-Z_{r_1,r_2,s_1,s_2}$, the projection of S_z on the xy-plane intersects the rectangle

$$\{(x, y, 0): r_1<x<r_2,\ s_1<y<s_2,\ x\in\mathscr{E},\ y\in\mathscr{E}\}$$

in a planar set of power $2^{\aleph_0}$, second category, and positive exterior measure. Set $Z^*=\bigcup Z_{r_1,r_2,s_1,s_2}$, where r_1, r_2, s_1, s_2 range over the set of rational numbers so that $r_1<r_2$ and $s_1<s_2$. Then Z^* is a subset of $\mathscr{E}$ of first category, and for every $z_0\in\mathscr{E}-Z^*$, the z-plane $z=z_0$ intersects E_3 in a planar set that is everywhere of power $2^{\aleph_0}$, everywhere of second category, and everywhere of positive exterior measure.

Consider the following propositions:

($\mathfrak{U}_K^*$) *$\mathscr{S}=E_1\cup E_2\cup E_3$, where every x-plane intersects E_1 in at most $\aleph_0$ points, every y-plane intersects E_2 in at most $\aleph_0$ points, and every member of some non-enumerable set of z-planes intersects E_3 in a planar set of first category.*

($\mathfrak{U}_M^*$) *$\mathscr{S}=E_1\cup E_2\cup E_3$, where every x-plane intersects E_1 in at most $\aleph_0$ points, every y-plane intersects E_2 in at most $\aleph_0$ points, and every member of some non-enumerable set of z-planes intersects E_3 in a planar set of measure zero.*

($\mathfrak{N}^*$) *$\mathscr{E}$ is not the union of less than $2^{\aleph_0}$ linear sets of first category.*

($\mathfrak{M}^*$) *$\mathscr{E}$ is not the union of less than $2^{\aleph_0}$ linear sets of measure zero.*

We shall show that they are related to the continuum hypothesis

(H) $2^{\aleph_0}=\aleph_1$.

THEOREM 3. *The conjunction of* ($\mathfrak{N}^*$) *and* ($\mathfrak{U}_K^*$) *is equivalent to* (H).

PROOF. (a) Clearly (H) implies ($\mathfrak{N}^*$). According to SIERPIŃSKI [4, p. 11, Theorem 8], (H) implies that $\mathscr{S}=E_1\cup E_2\cup E_3$, where every x- (y-, z-) plane intersects E_1 (E_2, E_3) in at most $\aleph_0$ points, and this evidently implies ($\mathfrak{U}_K^*$).

(b) Assume that $(\mathfrak{R}^*)$ and $(\mathfrak{U}_K^*)$ are true, but that (H) is false. Then, because of the hypothesis concerning E_3, there exists a cardinal number $\mathfrak{p}$ with $\aleph_0 < \mathfrak{p} < 2^{\aleph_0}$, and a set P of $\mathfrak{p}$ real numbers, such that, for every $c \in P$, the set

$$M_c = \{(x, y, c):\ x \in \mathscr{E},\ y \in \mathscr{E}\} \cap E_3$$

is a planar set of first category. Hence, by a theorem of KURATOWSKI and ULAM [3], there exists a subset F_c of $\mathscr{E}$ of first category such that, for every $b \in \mathscr{E} - F_c$, the set

$$\{(x, b, c):\ x \in \mathscr{E}\} \cap E_3$$

is a linear set of first category. Let $F = \bigcup_{c \in P} F_c$. Then $(\mathfrak{R}^*)$ implies that the set $\mathscr{E} - F$ is not empty. If $b_0 \in \mathscr{E} - F$, then, for every $c \in P$, the set

$$H_c = \{(x, b_0, c):\ x \in \mathscr{E}\} \cap E_3$$

is a linear set of first category, and $(\mathfrak{R}^*)$ implies that the projection of the set $H = \bigcup_{c \in P} H_c$ on the x-axis does not exhaust the axis. Thus, there exists a number $a_0 \in \mathscr{E}$ such that, for every $c \in P$, the point (a_0, b_0, c) does not belong to E_3 and consequently does belong to $E_1 \cup E_2$. But there are $\mathfrak{p}$ such points, all of which lie on a straight line parallel to the z-axis. Therefore either $\mathfrak{p}$ of these points are in E_1, in which case the x-plane $x = a_0$ contains $\mathfrak{p}$ points of E_1; or $\mathfrak{p}$ of these points are in E_2, in which case the y-plane $y = b_0$ contains $\mathfrak{p}$ points of E_2. But this contradicts $(\mathfrak{U}_K^*)$, and so (H) must be true.

THEOREM 4. *The conjunction of* $(\mathfrak{M}^*)$ *and* $(\mathfrak{U}_M^*)$ *is equivalent to* (H).

PROOF. In the proof of Theorem 3, replace "$(\mathfrak{R}^*)$" by "$(\mathfrak{M}^*)$", "$(\mathfrak{U}_K^*)$" by "$(\mathfrak{U}_M^*)$", "first category" by "measure zero", and "KURATOWSKI and ULAM [3]" by "FUBINI [2]".

The following propositions, the second pair of which were introduced in [1], are also related to (H):

$(\mathfrak{U}_K)$ *$\mathscr{S} = E_1 \cup E_2 \cup E_3$, where every member of some set of second category of x-planes intersects E_1 in at most $\aleph_0$ points, every member of some set of second category of y-planes intersects E_2 in less than $2^{\aleph_0}$ points, and every member of some nonenumerable set of z-planes intersects E_3 in a planar set of first category.*

$(\mathfrak{U}_M)$ *$\mathscr{S} = E_1 \cup E_2 \cup E_3$, where every member of some set of positive exterior measure of x-planes intersects E_1 in at most $\aleph_0$ points, every member of some set of positive exterior measure of y-planes intersects E_2 in less than $2^{\aleph_0}$ points, and every member of some nonenumerable set of z-planes intersects E_3 in a planar set of measure zero.*

$(\mathfrak{R})$ *The union of less than $2^{\aleph_0}$ linear sets of first category is of first category.*

$(\mathfrak{M})$ *The union of less than $2^{\aleph_0}$ linear sets of measure zero is of measure zero.*

THEOREM 5. *The conjunction of* $(\mathfrak{R})$ *and* $(\mathfrak{U}_K)$ *is equivalent to* (H).

PROOF. (a) To show that (H) implies $(\mathfrak{R})$ and $(\mathfrak{U}_K)$, replace "$(\mathfrak{R}^*)$" by "$(\mathfrak{R})$" and "$(\mathfrak{U}_K^*)$" by "$(\mathfrak{U}_K)$" in part (a) of the proof of Theorem 3.

(b) Assume that $(\mathfrak{R})$ and $(\mathfrak{U}_K)$ are true, but that (H) is false. Then, because of the hypothesis concerning E_3, there exists a cardinal number $\mathfrak{p}$ with $\aleph_0 < \mathfrak{p} < 2^{\aleph_0}$,

and a set P of $\mathfrak{p}$ real numbers, such that, for every $c \in P$, the set

$$M_c = \{(x, y, c): x \in \mathscr{E},\ y \in \mathscr{E}\} \cap E_3$$

is a planar set of first category. Hence, according to a theorem of KURATOWSKI and ULAM [3], there exists a subset F_c of $\mathscr{E}$ of first category such that, for every $b \in \mathscr{E} - F_c$, the set

$$\{(x, b, c): x \in \mathscr{E}\} \cap E_3$$

is a linear set of first category. Let $F = \bigcup_{c \in P} F_c$. Then $(\mathfrak{R})$ implies that F is a subset of $\mathscr{E}$ of first category. Because of the hypothesis concerning E_2, if we put $Y = \mathscr{E} - F$, there exists a number $b_0 \in Y$ such that the y-plane $y = b_0$ intersects E_2 in less than $2^{\aleph_0}$ points, and, for every $c \in P$, the set

$$H_c = \{(x, b_0, c): x \in \mathscr{E}\} \cap E_3$$

is a linear set of first category. Now $(\mathfrak{R})$ implies that the projection of the set $H = \bigcup_{c \in P} H_c$ on the x-axis is a linear set of first category. Thus, there exists a subset X of $\mathscr{E}$ such that $\mathscr{E} - X$ is a linear set of first category and, for every $a_0 \in X$ and every $c \in P$, the point (a_0, b_0, c) does not belong to E_3 and consequently does belong to $E_1 \cup E_2$. Because of the hypothesis concerning E_1, it follows from $(\mathfrak{R})$ that there exists a subset X^* of X with $|X^*| = 2^{\aleph_0}$ such that, for every $a_0 \in X^*$, $\mathfrak{p}$ of the points (a_0, b_0, c) with $c \in P$ belong to E_2. This implies that the y-plane $y = b_0$ contains $2^{\aleph_0} \cdot \mathfrak{p} = 2^{\aleph_0}$ points of E_2, contrary to the choice of b_0. Hence, (H) must be true if $(\mathfrak{R})$ and $(\mathfrak{U}_K)$ are true.

THEOREM 6. *The conjunction of* $(\mathfrak{M})$ *and* $(\mathfrak{U}_M)$ *is equivalent to* (H).

PROOF. (a) To show that (H) implies $(\mathfrak{M})$ and $(\mathfrak{U}_M)$, replace "$(\mathfrak{R}^*)$" by "$(\mathfrak{M})$" and "$(\mathfrak{U}_K^*)$" by "$(\mathfrak{U}_M)$" in part (a) of the proof of Theorem 3.

(b) To prove that the conjunction of $(\mathfrak{M})$ and $(\mathfrak{U}_M)$ implies (H), replace "$(\mathfrak{R})$" by "$(\mathfrak{M})$", "$(\mathfrak{U}_K)$" by "$(\mathfrak{U}_M)$", "first category" by "measure zero", and "KURATOWSKI and ULAM [3]" by "FUBINI [2]" in part (b) of the proof of Theorem 5.

[1] BAGEMIHL, F.: Some results connected with the continuum hypothesis. Z. Math. Logik Grundlagen Math. **5**, 97–116 (1959). — [2] FUBINI, G.: Sugli integrali multipli. Atti Accad. Naz. Lincei, Rend. (5) $\mathbf{16}_{\mathrm{I}}$, 608–614 (1907). — [3] KURATOWSKI, C., and S. ULAM: Quelques propriétés topologiques du produit combinatoire. Fund. Math. **19**, 247–251 (1932). — [4] SIERPIŃSKI, W.: Sur quelques propositions concernant la puissance du continu. Fund. Math. **38**, 1–13 (1951).

University of Notre Dame, Notre Dame, Ind., U.S.A.

(Eingegangen am 27. April 1959)

Kasch, F., u. B. Volkmann
Math. Zeitschr. 72, 367–378 (1960)

Metrische Sätze über transzendente Zahlen in P-adischen Körpern

Von
Friedrich Kasch und Bodo Volkmann

1. Fragestellung. Die auf Mahler zurückgehende Klasseneinteilung der transzendenten Zahlen[1]) ist von ihm [6] auch auf Zahlen in den P-adischen Körpern K_P übertragen worden. 1936 führte dann Turkstra [8] eine Maßtheorie in den Körpern K_P ein, die auch die Übertragung einiger damals im Reellen bzw. Komplexen bekannter metrischer Aussagen auf das P-adische gestattete. Insbesondere ergab sich als Analogon zu einem Satz von Mahler die Aussage, daß im Sinne des Turkstraschen Maßes fast alle P-adischen Zahlen S-Zahlen sind. Dabei bewies Turkstra als Gegenstück zu dem Ergebnis von Mahler, wonach fast alle reellen Zahlen ξ S-Zahlen von einem Typ[2])

$$\vartheta(\xi) \leqq 4 \tag{1}$$

sind, daß für fast alle P-adischen Zahlen die Ungleichung

$$\vartheta(\xi) \leqq 7 \tag{2}$$

erfüllt ist.

Seither ist die Abschätzung (1) mehrfach verbessert worden, und das kürzlich von uns erzielte, bisher beste Ergebnis [3] besagt insbesondere, daß fast alle reellen ξ den Ungleichungen

$$1 \leqq \vartheta_n(\xi) \leqq 2 - \frac{2}{n} \qquad (n = 3, 4, \ldots) \tag{3}$$

genügen. Dagegen liegen im P-adischen Fall, soweit den Verfassern bekannt, bisher noch keine Resultate vor, die über (2) hinausgehen.

In der vorliegenden Arbeit wird für eine beliebige Primzahl P folgende Verschärfung von (2) bewiesen, die ungefähr die Güte von (3) erreicht:

Satz 1. *Für fast alle Zahlen $\xi \in K_P$ gilt*[3])

$$\vartheta_1(\xi) = 2, \tag{4}$$

$$1 + \frac{1}{n} \leqq \vartheta_n(\xi) \leqq 2 - \frac{1}{2n} \qquad (n = 2, 3, \ldots). \tag{5}$$

[1]) Siehe z. B. Schneider [7].
[2]) Siehe [3], S. 442.
[3]) Die genauen Definitionen der hier benutzten Begriffe sind im nächsten Abschnitt enthalten.

Der Beweis erfolgt durch vollständige Induktion nach n mit dem Induktionsbeginn $n=1$. Sonst folgen wir methodisch im Wesentlichen unserem Beweis von (3), dessen Grundgedanke darin besteht, Eigenschaften der Diskriminante der approximierenden Polynome für die Abschätzung des Maßes auszunutzen.

Im Fall $n=2$ hat für reelle Zahlen KUBILYUS [4] das bestmögliche Ergebnis, nämlich

$$\vartheta_2(\xi) = 1 \quad \text{für fast alle } \xi,$$

erzielt. Sein Beweis wurde von einem der Verfasser [2] durch einen einfacheren ersetzt, wobei sich zugleich das bestmögliche Ergebnis für die komplexen Zahlen mit ergab. Durch Übertragung dieser Methode erhalten wir auch im P-adischen Fall die bestmögliche Aussage:

Satz 2. *Für fast alle Zahlen $\xi \in K_P$ gilt*

$$\vartheta_2(\xi) = \tfrac{3}{2}. \tag{6}$$

2. Voraussetzungen. Es sei P eine (für das Folgende feste) Primzahl und K_P der zugehörige Henselsche Körper. Die Bewertung $|\ |_P$ von K_P läßt sich dann eindeutig auf die algebraisch abgeschlossene Hülle $\hat{K}_P$ von K_P fortsetzen und soll auch in diesem Sinne mit $|\ |_P$ bezeichnet werden. Alle im folgenden auftretenden algebraischen Zahlen sind als Elemente von $\hat{K}_P$ zu betrachten.

Wenn n und H natürliche Zahlen sind, sei $\mathfrak{Q}(n, H)$ die Menge der Polynome $q(x)$ mit ganzrationalen Koeffizienten a_i von einem Grad $\leqq n$ und einer Höhe $(= \max |a_i|) \leqq H$. Dann setzen wir für[4]) jedes $\xi \in K_P$

$$w_n(H, \xi) = \min_{\substack{q \in \mathfrak{Q}(n,H) \\ q(\xi) \neq 0}} |q(\xi)|_P,$$

ferner

$$w_n(\xi) = \overline{\lim_{H \to \infty}} \frac{\log \frac{1}{w_n(H, \xi)}}{\log H} \quad \text{sowie} \quad \vartheta_n(\xi) = \frac{w_n(\xi)}{n},$$

schließlich

$$\vartheta(\xi) = \sup_{n=1,2,\ldots} \vartheta_n(\xi) \quad \text{und} \quad w(\xi) = \overline{\lim_{n \to \infty}} \vartheta_n(\xi).$$

Eine Zahl $\xi \in K_P$ heißt *(P-adische) S-Zahl*, wenn $0 < \vartheta(\xi) < \infty$ ist.

Das Turkstrasche Maß wird in Anlehnung an die von FELLER und TORNIER [1] für den Baireschen Nullraum entwickelte Maßtheorie folgendermaßen eingeführt[5]): Man definiert als *P-adisches Intervall n-ter Stufe* (bei ganzrationalem n) die Menge aller Zahlen $\xi \in K_P$, in deren P-adischer Entwicklung

$$\xi = \sum_{\nu=-\infty}^{\infty} a_\nu P^\nu \qquad \left(0 \leqq a_\nu < P;\ a_\nu \text{ ganzrational};\ a_\nu = 0 \quad \text{für} \quad \nu < \nu_0(\xi)\right)$$

alle Koeffizienten a_ν mit $\nu < n$ denen einer vorgegebenen Zahl gleich sind. Es zeigt sich, daß jedes P-adische Intervall sowohl offen als auch abgeschlossen ist, und ferner läßt sich beweisen, daß jede offene Menge $O \subseteqq K_P$ als Ver-

[4]) Vgl. [7] sowie [3].

[5]) Siehe [8], Kap. V.

einigung von P-adischen Intervallen darstellbar ist. Als *beschränkt* wird eine Menge $M \subseteqq K_P$ bezeichnet, wenn es eine ganzrationale Zahl t so gibt, daß $M \subseteqq K_P^t$ (= Menge der $\xi \in K_P$ mit $|\xi|_P \leqq P^t$) gilt.

Der Aufbau der Maßtheorie wird dann in mehreren Schritten vollzogen:

1. Für jedes P-adische Intervall E von n-ter Stufe setzt man

$$T(E) = \sup_{\alpha, \beta \in E} |\alpha - \beta|_P = P^{-n}.$$

2. Für jede beschränkte, offen-abgeschlossene Menge $A \subseteqq K_P$ wird, wenn $A = \bigcup_i E^i$ eine Zerlegung in paarweise fremde P-adische Intervalle ist,

$$T(A) = \sum_i T(E^i)$$

gesetzt und gezeigt, daß die linke Seite nicht von der speziellen Wahl der Zerlegung abhängt.

3. Für jede beschränkte, offene Menge $O \subseteqq K_P$ wird eine Zerlegung $O = \bigcup_i E^i$ in *maximale P-adische Intervalle* (d.h. solche, die in keinem ganz zu O gehörenden P-adischen Intervall echt enthalten sind) eingeführt, die Eindeutigkeit dieser Zerlegung bewiesen und

$$T(O) = \sum_i T(E^i)$$

gesetzt. Daraus erhält man Analoga zu den bekannten Sätzen der Inhaltstheorie für lineare, offene Mengen.

4. Für jede beliebige, beschränkte Menge $M \subseteqq K_P$ wird das äußere Turkstrasche Maß

$$\overline{T}(M) = \inf T(O)$$

eingeführt, wobei O alle offenen, beschränkten Mengen mit $M \subseteqq O$ durchläuft. Es folgt insbesondere, daß für beschränkte, offene Mengen O stets $\overline{T}(O) = T(O)$ ist. Sodann lassen sich in bekannter Weise das zugehörige innere Maß und der Begriff der Meßbarkeit definieren.

5. Für jede beliebige Menge $M \subseteqq K_P$ wird $M \cap K_P^t = M^t$ und

$$\overline{T}(M) = \lim_{t \to +\infty} \overline{T}(M^t)$$

gesetzt. Damit wird auch für nichtbeschränkte Mengen M der Begriff der Meßbarkeit und des Maßes $T(M)$ erklärt, sofern dieses existiert. Man kann dann u. a. nachweisen, daß das Turkstrasche Maß $T(M)$ monoton und σ-additiv ist.

3. Hilfssätze. Zu den Beweisen benötigen wir die folgenden Hilfssätze[6]:

Hilfssatz 1. *Es seien*

$$L_i(x_0, \ldots, x_n) = L_i(\mathfrak{x}) = \sum_{j=0}^{n} \gamma_{ij} x_j \qquad (i = 0, \ldots, n)$$

[6]) Dabei sind $C_1, C_2, \ldots$ positive, reelle Zahlen, die von n und P, nicht aber von H abhängen dürfen. Alle im folgenden auftretenden Polynome haben ganzrationale Koeffizienten.

$n+1$ Linearformen mit reellen Koeffizienten und einer nichtverschwindenden Determinante d; ferner sei

$$L(\mathfrak{x}) = \sum_{j=0}^{n} \gamma_j x_j$$

eine Linearform mit P-adisch ganzen Koeffizienten. Sind dann $\beta, \beta_0, \ldots, \beta_n$ positive reelle Zahlen mit $0<\beta<1$ und $\beta\beta_0 \ldots \beta_n \geqq |d| P$, so gibt es einen Gitterpunkt $\mathfrak{x}$ mit $|L(\mathfrak{x})|_P<\beta$ und

$$|L_i(\mathfrak{x})| \leqq \beta_i \qquad (i = 0, \ldots, n).$$

Beweis. Folgt aus [5], Satz 1.

Hilfssatz 2. *Zu jeder Linearform $L(\mathfrak{x})=\gamma_0 x_0+\cdots+\gamma_n x_n$ mit P-adisch ganzen Koeffizienten gibt es für jedes hinreichend große $t>1$ wenigstens einen Gitterpunkt $\mathfrak{x} \neq (0, \ldots, 0)$ mit $\max(|x_0|, \ldots, |x_n|)=H$, derart, daß $H \leqq t^{1/n}$ und $|L(\mathfrak{x})|_P<\frac{P}{t^{1+1/n}}$ ist.*

Beweis[7]. In dem Ungleichungssystem

$$\begin{aligned} |\gamma_0 x_0+\gamma_1 x_1+\cdots+\gamma_n x_n|_P &< \frac{P}{t^{1+1/n}} \\ |x_0| &\leqq t^{1/n} \\ |x_1| &\leqq t^{1/n} \\ \ddots &\quad \vdots \\ |x_n| &\leqq t^{1/n} \end{aligned}$$

ist die Determinante der letzten $n+1$ Ungleichungen offenbar gleich 1. Daher sind die Voraussetzungen von Hilfssatz 1 mit $\beta=\frac{P}{t^{1+1/n}}$ und $\beta_0=\beta_1=\cdots=t^{-1/n}$ erfüllt, und somit folgt die Behauptung.

Hilfssatz 3. *Für jede transzendente, insbesondere also für jede S-Zahl $\xi \in K_P$ ist*

$$\vartheta_n(\xi) \geqq 1+\frac{1}{n} \qquad (n = 1, 2, \ldots).$$

Beweis[8]. Für jede P-adisch ganze Zahl ξ und jedes hinreichend große t ergibt sich durch Anwendung von Hilfssatz 2 mit $\gamma_i=\xi^i$ und $x_i=a_i$ $(i=0, \ldots, n)$ die Existenz eines Polynoms $q(x) \in \mathfrak{Q}(n, [t^{1/n}])$ mit $|q(\xi)|_P \leqq P t^{-(1+1/n)}$. Folglich ist

$$w_n(H, \xi) < \frac{C_1}{H^{n+1}} \qquad (H = 1, 2, \ldots),$$

woraus die Behauptung leicht zu folgern ist.

Hilfssatz 4. *Sei $q(x)=a_0 x^n+a_1 x^{n-1}+\cdots+a_n$ ein Polynom mit den Nullstellen $\alpha_1, \alpha_2, \ldots, \alpha_n$. Dann gilt, wenn die Indizes $i_1, \ldots, i_k$ voneinander verschieden sind, die Ungleichung*

$$|a_0 \alpha_{i_1} \ldots \alpha_{i_k}|_P \leqq 1.$$

[7]) Siehe [8], Kap. V, Satz 2*.

[8]) Vgl. [8], S. 67–68.

Beweis. Nach dem Beweis von [7], Hilfssatz 17 ist jedes Produkt der Form $a_0\alpha_{i_1}\dots\alpha_{i_k}$ ganzalgebraisch, und daraus folgt die Behauptung.

Hilfssatz 5. *Sei $q(x)$ ein Polynom vom Grade $n\geqq 2$ mit lauter verschiedenen Nullstellen $\alpha_1,\dots,\alpha_n$. Dann gilt, wenn $D(q)$ die Diskriminante bezeichnet,*

$$|q'(\alpha_m)|_P\geqq\sqrt{|D(q)|_P}\qquad(m=1,\dots,n).$$

Beweis. Nach [3], Gl. (4) ist

$$|q'(\alpha_m)|_P=\frac{\sqrt{|D(q)|_P}}{\left|a_0^{n-2}\prod\limits_{i,j\neq m;i<j}(\alpha_i-\alpha_j)\right|_P}.$$

Bei der Ausmultiplikation des Produktes im Nenner entsteht eine Summe von Produkten mit je $\binom{n-1}{2}$ Faktoren aus der Menge der $n-1$ Wurzeln α_i mit $i\neq m$. Da in diesen Produkten jedes α_i höchstens in der $(n-2)$-ten Potenz vorkommt, läßt sich jedes von ihnen in Faktoren der Form $a_0\alpha_{i_1}\dots\alpha_{i_k}$ zerlegen. Somit hat nach Hilfssatz 4 jeder Summand des Nenners einen P-adischen Wert $\leqq 1$, so daß

$$\left|a_0^{n-2}\prod_{i,j\neq m;i<j}(\alpha_i-\alpha_j)\right|_P\leqq 1$$

und damit die Behauptung folgt.

Hilfssatz 6. *Unter den Voraussetzungen von Hilfssatz 5 gilt, wenn ξ eine beliebige, von $\alpha_1,\dots,\alpha_n$ verschiedene Zahl aus K_P mit*

$$|\xi-\alpha_m|_P=\min_{i=1,\dots,n}|\xi-\alpha_i|_P$$

ist,

$$|\xi-\alpha_m|_P\leqq\frac{|q(\xi)|_P}{\sqrt{|D(q)|_P}}.$$

Beweis. Für jedes $i\neq m$ ist

$$|\alpha_m-\alpha_i|_P=|(\alpha_m-\xi)+(\xi-\alpha_i)|_P\leqq|\alpha_i-\xi|_P,$$

und somit wird

$$|q'(\alpha_m)|_P=|a_0|_P\prod_{\substack{i=1\\i\neq m}}^{n}|\alpha_m-\alpha_i|_P\leqq|a_0|_P\prod_{\substack{i=1\\i\neq m}}^{n}|\alpha_i-\xi|_P.$$

Folglich ergibt sich nach Hilfssatz 5

$$\sqrt{|D(q)|_P}\leqq|a_0|_P\prod_{\substack{i=1\\i\neq m}}^{n}|\alpha_i-\xi|_P$$

und damit

$$|\xi-\alpha_m|_P\leqq\frac{|a_0|_P}{\sqrt{|D(q)|_P}}\prod_{i=1}^{n}|\xi-\alpha_i|_P=\frac{|q(\xi)|_P}{\sqrt{|D(q)|_P}}.$$

Hilfssatz 7. *Seien $q_0(x)$, $q_1(x),\dots,q_r(x)$ Polynome mit*

$$q_0(x)=\prod_{i=1}^{r}q_i(x),\qquad\text{Grad } q_0=n \text{ und Höhe } q_i=H_i\qquad(i=0,\dots,n).$$

Dann gilt $H_0\geqq(4n)^{-n}\prod\limits_{i=1}^{r}H_i$.

Beweis. Siehe [7], Hilfssatz 16.

Hilfssatz 8. *Für jedes Polynom $q(x)$ vom Grad n gilt wenigstens eine der folgenden Aussagen*[9]*:*

1. q ist Produkt von n linearen Polynomen,

2. q ist Produkt von zwei Polynomen mindestens zweiten Grades,

3. q hat lauter verschiedene Nullstellen.

Beweis. Im Ring der Polynome mit ganzrationalen Koeffizienten sei

$$q = q_1^{e_1} \cdots q_r^{e_r}$$

die eindeutige Zerlegung von q in Primelementpotenzen; ferner sei Grad $q_i = n_i$ $(i = 1, \ldots, r)$. Wird nun angenommen, es gelte weder 1. noch 3., so gibt es wenigstens ein n_i, etwa n_1, mit $1 < n_i < n$. Dann kann auch nicht $n_1 = n - 1$ sein, weil daraus 3. bzw. 1. folgen würde. Somit gilt $1 < n_1 < n - 1$, woraus unmittelbar folgt, daß

$$q = q_1 (q_1^{e_1 - 1} q_2^{e_2} \cdots q_r^{e_r})$$

eine Zerlegung mit der Eigenschaft 2. ist.

Hilfssatz 9. *Bei vorgegebenen ganzrationalen Zahlen $H > 1$ und x mit $H^2 \neq x$ hat die Gleichung*

$$H^2 - 4 a_0 a_1 = x$$

höchstens $C_2 H^{C_3/\log\log H}$ Lösungen in ganzen Zahlen a_0, a_1 mit $|a_0| \leqq H$, $|a_1| \leqq H$.

Beweis[10]. Ist $d(M)$ die Anzahl der positiven Teiler von M, so beträgt die Lösungsanzahl höchstens

$$2 d(H^2 - x) \leqq C_2 H^{C_3/\log\log H}.$$

Hilfssatz 10. *Bei vorgegebenen ganzrationalen Zahlen $H > 1$ und x hat die Gleichung*

$$a_1 - 4 H a = x$$

höchstens $C_2 r(x, 4H) H^{C_3/\log\log H}$ Lösungen in ganzen Zahlen a_1, a mit $|a_1| \leqq H$, $|a| \leqq H$, wobei $r(u, v)$ der größte gemeinsame quadratische Teiler von u und v ist.

Beweis. Siehe [2], Seite 267.

Hilfssatz 11. *Sind N und M natürliche Zahlen, so gilt*

$$\sum_{x=1}^{M} r(x, N) \leqq M \log N.$$

Beweis. Sei $d^2 | N$; dann gilt $r(x, N) = d$ genau für die x von der Form $x = d^2 x_1$ mit $r(x_1, N) = 1$. Es gibt also höchstens M/d^2 Zahlen x mit dieser Eigenschaft, und daher wird

$$\sum_{x=1}^{M} r(x, N) \leqq \sum_{d^2 | N} \frac{M}{d^2}\, d \leqq M \sum_{d | N} \frac{1}{d} \leqq C_4 M \log N.$$

4. Beweis von Satz 1. Im Hinblick auf Hilfssatz 3 ist nur noch

$$\vartheta_1(\xi) \leqq 2 \quad \text{und} \quad \vartheta_n(\xi) \leqq 2 - \frac{1}{2n} \qquad (n = 2, 3, \ldots) \tag{7}$$

[9]) Dabei ist an die oben getroffene Vereinbarung zu erinnern, wonach nur Polynome über dem Ring der ganzrationalen Zahlen betrachtet werden.

[10]) Vgl. [2], S. 266.

für fast alle $\xi \in K_p$ zu beweisen. Da die Menge der algebraischen Zahlen aus K_P das Maß 0 besitzt, können wir uns dabei auf die Menge K der transzendenten P-adischen Zahlen beschränken.

Sei $S_n(\sigma)$ die Menge der $\xi \in K$ mit $\vartheta_n(\xi) > \sigma$. Dann gibt es, wie man der Definition leicht entnimmt, zu jeder Zahl $\xi \in S_n(\sigma)$ unendlich viele H, zu denen Polynome $q(x) \in \mathfrak{Q}(n, H)$, $q(x) \notin \mathfrak{Q}(n, H-1)$ mit

$$|q(\xi)|_P < H^{-n\sigma} \tag{8}$$

existieren; die Menge aller $\xi \in K$, für die dies zutrifft, bildet also eine Obermenge $S_n^*(\sigma)$ von $S_n(\sigma)$. Ist nun für alle $\sigma > \sigma_0$ die Gleichung $T(S_n^*(\sigma)) = 0$ bewiesen, so folgt auch $T(S_n^*(\sigma_0)) = T(S_n(\sigma_0)) = 0$; denn $S_n^*(\sigma_0)$ ist als Vereinigung der abzählbar vielen Mengen $S_n^*\left(\sigma_0 + \frac{1}{j}\right)$ $(j = 1, 2, \ldots)$ darstellbar. Wir werden diese Überlegung im Fall $n = 1$ mit $\sigma_0 = 2$ und im Fall $n \geqq 2$ mit $\sigma_0 = 2 - \frac{1}{2n}$ anwenden, so daß dann (7) folgt.

Um für jedes $\varepsilon > 0$ die Gleichung $T(S_1^*(2+\varepsilon)) = 0$ zu beweisen, bemerken wir, daß in diesem Fall für jede Lösung von (8) wegen[11])

$$|q(\xi)|_P = |a_0 \xi + a_1|_P = |a_0|_P \cdot \left|\xi + \frac{a_1}{a_0}\right|_P < H^{-2-\varepsilon}$$

offenbar

$$\left|\xi + \frac{a_1}{a_0}\right|_P < \frac{1}{|a_0|_P H^{2+\varepsilon}} = \frac{P^{\varkappa}}{H^{2+\varepsilon}}$$

gilt, wenn $|a_0|_P = P^{-\varkappa}$ ist. Wir überdecken nun jede rationale Zahl $-\frac{a_1}{a_0}$ durch das P-adische Intervall $\mathfrak{i}(a_0, a_1)$ der Zahlen $\xi \in K_P$ mit $\left|\xi + \frac{a_1}{a_0}\right|_P < \frac{H^{-2-\varepsilon}}{|a_0|_P}$. Dann ist offenbar die Menge $S_1^*(2+\varepsilon)$ für jedes H_0 in der Vereinigung

$$\bigcup_{H=H_0}^{\infty} \bigcup_{\max(|a_0|, |a_1|)=H} \mathfrak{i}(a_0, a_1)$$

enthalten. Für das Turkstrasche Maß verwenden wir die Zerlegung

$$T_H = T\Big(\bigcup_{\max(|a_0|, |a_1|)=H} \mathfrak{i}(a_0, a_1)\Big) \leqq T\Big(\bigcup_{|a_0| \leqq H = |a_1|} \mathfrak{i}(a_0, a_1)\Big) + T\Big(\bigcup_{|a_1| < H = |a_0|} \mathfrak{i}(a_0, a_1)\Big),$$

und die beiden letzten Summanden nennen wir T_H' bzw. T_H''.

Zur Abschätzung von T_H' ist zu berücksichtigen, daß bei den auftretenden Intervallen stets $0 \leqq \varkappa \leqq k = \left[\frac{\log H}{\log P}\right]$ ist und jedes solche $\varkappa$ höchstens $2H P^{-\varkappa}$ mal vorkommt. Daher wird

$$T_H' \leqq \sum_{\varkappa=0}^{k} \frac{2H}{P^{\varkappa}} \cdot \frac{P^{\varkappa}}{H^{2+\varepsilon}} = 2(k+1) \frac{1}{H^{1+\varepsilon}} < \frac{C_5}{H^{1+\varepsilon/2}}.$$

Für T_H'' erhält man entsprechend die Ungleichung

$$T_H'' < \frac{C_6 H}{|H|_P H^{2+\varepsilon}} = \frac{C_6}{|H|_P H^{1+\varepsilon}},$$

[11]) Da keine konstanten Polynome als Lösungen auftreten können, haben alle zu betrachtenden $q(x)$ den genauen Grad 1.

und somit ergibt sich

$$T\left(S_1^*(2+\varepsilon)\right) \leqq \lim_{H_0\to\infty} \sum_{H=H_0}^{\infty} (T_H' + T_H'') \leqq \lim_{H_0\to\infty} \sum_{H=H_0}^{\infty} \frac{C_5}{H^{1+\varepsilon/2}} + \lim_{H_0\to\infty} \sum_{H=H_0}^{\infty} \frac{C_6}{|H|_P\, H^{1+\varepsilon}}.$$

Da der erste Grenzwert verschwindet, bleibt nur der zweite zu betrachten. Es ist[12])

$$\sum_{H=1}^{\infty} \frac{1}{|H|_P\, H^{1+\varepsilon}} = \sum_{l=0}^{\infty} \sum_{\substack{H=1 \\ P^l \mid H}}^{\infty} \frac{P^l}{H^{1+\varepsilon}} < \sum_{l=0}^{\infty} \sum_{m=1}^{\infty} \frac{P^l}{(m\,P^l)^{1+\varepsilon}} = \frac{P^\varepsilon}{P^\varepsilon - 1} \sum_{m=1}^{\infty} \frac{1}{m^{1+\varepsilon}} < \infty,$$

woraus

$$\lim_{H_0\to\infty} \sum_{H=H_0}^{\infty} T_H'' = 0$$

folgt, und daher gilt $T\left(S_1^*(2+\varepsilon)\right) = 0$.

Es sei nun bereits für $m = 1, 2, \ldots, n-1$ die Gleichung $T\left(S_m^*(\sigma_0)\right) = 0$ mit

$$\sigma_0 = \sigma_0(m) = \begin{cases} 2 & \text{für } m = 1, \\ 2 - \dfrac{1}{2m} & \text{für } m = 2, \ldots, n-1 \end{cases} \tag{9}$$

bewiesen. Dann ist zu zeigen, daß für jedes $\varepsilon > 0$, wenn $\sigma = 2 - \frac{1}{2n} + \varepsilon$ gesetzt wird, $T\left(S_n^*(\sigma)\right) = 0$ gilt. Dazu stellen wir die Menge $S = S_n^*(\sigma)$ als Vereinigung von vier Teilmengen S^0, S^1, S^2, S^3 dar, die folgendermaßen definiert sind: Es sei $S^0(\sigma)$ die Menge der $\xi \in K$, für die (8) unendlich viele Lösungen $q(x)$ mit Grad $q(x) < n$ besitzt; ferner seien $S^1(\sigma)$, $S^2(\sigma)$ und $S^3(\sigma)$ die Menge der $\xi \in K$, für die (8) unendlich viele Lösungen $q(x)$ besitzt, die vom Grad n sind und außerdem im Sinne von Hilfssatz 8 die Aussagen 1., 2. bzw. 3. erfüllen. Dann gilt sicher $S = S^0 \cup S^1 \cup S^2 \cup S^3$. Wir setzen nun $V = \bigcup_{m=1}^{n-1} S_m^*\left(\sigma_0(m)\right)$ und zeigen zunächst

$$T(S^0 \cup S^1 \cup S^2) \leqq T(V), \tag{10}$$

woraus $T(S^0 \cup S^1 \cup S^2) = 0$ folgt, da nach Induktionsvoraussetzung $T(V) = 0$ ist.

Ist $\xi \in S^0$, so gibt es offenbar ein $m < n$ derart, daß unendlich viele Lösungen von (8) existieren, die den Grad m haben. Wegen

$$n\left(2 - \frac{1}{2n} + \varepsilon\right) > m\,\sigma_0(m) \qquad (m = 1, \ldots, n-1) \tag{11}$$

folgt dann $\xi \in S_m^*\left(\sigma_0(m)\right)$, also auch $\xi \in V$ und damit $S^0 \subseteq V$.

Ist $\xi \in S^1$, so gibt es unendlich viele Lösungen der Form $q = q_1 \ldots q_n$ mit linearen Polynomen q_i von (8). Bezeichnet man deren Höhen mit H_i und nimmt man $\xi \notin S_1^*(2)$ an, so besitzt die Ungleichung

$$|q_i(\xi)|_P < H_i^{-2}$$

nur endlich viele Lösungen in linearen Polynomen $q_i(x)$, für die wegen $\xi \in K$ stets $|q_i(\xi)|_P > 0$ gilt. Es gibt also eine Konstante $\gamma = \gamma(\xi) > 0$ mit

$$|q_i(\xi)|_P > \gamma\, H_i^{-2}$$

[12]) $P^l \mid H$ bedeutet, daß P^l, nicht aber P^{l+1} Teiler von H ist.

für alle linearen Polynome der Höhe H_i. Damit folgt unter Berücksichtigung von Hilfssatz 7

$$|q(\xi)|_P > \gamma'' \prod_{i=1}^{n} H_i^{-2} > C_7(\xi)\, H^{-2}.$$

Wegen (11) steht dies im Widerspruch zu der Voraussetzung, daß (8) unendlich viele Lösungen $q(x)$ der angegebenen Art hat. Folglich muß $S^1 \subseteq V \cup S_1^*(2)$ gelten, also auch $T(S^1) \leqq T(V)$.

Ist $\xi \in S^2$, so gibt es für wenigstens ein m unendlich viele Lösungen $q = q_1 q_2$ von (8) mit Grad $q = n$ und $1 < \text{Grad}\, q_1 = m < n$. Nimmt man $\xi \notin S_m^*(\sigma_0(m)) \cup S_{n-m}^*(\sigma_0(n-m))$ an, so folgt analog zur vorhergehenden Überlegung

$$|q(\xi)|_P > \gamma_1 \gamma_2 H_1^{-m\sigma_0(m)} H_2^{-(n-m)\sigma_0(n-m)} > C_8 H^{-\mu\sigma_0(\mu)}$$

mit $\mu = \min\big(m\sigma_0(m), (n-m)\sigma_0(n-m)\big)$. Wegen

$$n\left(2 - \frac{1}{2n} + \varepsilon\right) > \mu\sigma_0(\mu)$$

steht dies im Widerspruch zur Voraussetzung, daß (8) unendlich viele derartige Lösungen $q(x)$ besitzt. Folglich muß auch $T(S^2) \leqq T(V)$ gelten.

Ist $\xi \in S^3$, so gibt es unendlich viele Polynome $q(x)$ vom Grad n mit lauter verschiedenen Nullstellen, die (8) genügen. Wir ordnen jeder Nullstelle α_i eines solchen Polynoms das P-adische Intervall $\mathfrak{i}(\alpha_i)$ der Zahlen $\xi \in K_P$ mit

$$|\xi - \alpha_i|_P < \frac{1}{H^{n\sigma}\sqrt{|D(q)|_P}}$$

zu, wobei die linke Seite im Sinne der Bewertung in $\hat{K}_P$ zu verstehen ist. Dann gilt offenbar für alle diese Polynome in der Bezeichnungsweise von Hilfssatz 6 mit $\sigma = 2 - \frac{1}{2n} + \varepsilon$ die Ungleichung

$$|\xi - \alpha_m|_P \leqq \frac{|q(\xi)|_P}{\sqrt{|D(q)|_P}} < \frac{H^{-n\sigma}}{\sqrt{|D(q)|_P}},$$

d. h., es ist $\xi \in \mathfrak{i}(\alpha_m)$. Folglich wird für jedes H_0 die Menge S^3 durch die Gesamtheit dieser Intervalle mit $H \geqq H_0$ überdeckt, und somit gilt für das Turkstrasche Maß

$$\overline{T}(S^3) \leqq \lim_{H_0 \to \infty} \sum_{H=H_0}^{\infty} {\sum}' \frac{H^{-n\sigma}}{\sqrt{|D(q)|_P}}, \tag{12}$$

wobei in der inneren Summe $q(x)$ alle Polynome vom Grad n, der Höhe H und mit der (durch den Strich am Summenzeichen ausgedrückten) Nebenbedingung $D(q) \neq 0$ durchläuft.

Zur Abschätzung von

$$\Sigma = {\sum}' \frac{1}{\sqrt{|D(q)|_P}} \tag{13}$$

berücksichtigen wir, daß die Anzahl der Summanden $\leqq C_9 H^n$ ist, da jeweils einer der Koeffizienten, etwa a_i, den Betrag H hat. Wenn dann alle übrigen Koeffizienten bis auf einen, etwa a_j $(j \neq i)$, festgehalten werden, ist $D(q)$ ein Polynom $V(a_j)$ von einem Grad $g' \geqq 1$, so daß die Funktion $|V(a_j)|$ jeden

Wert höchstens $2g' = g$ mal annehmen kann. Die Summe Σ ist also (bei festem j und i) in eine Anzahl $\leqq C_{10} H^{n-1}$ von Summanden der Form

$$\Sigma^* = \sum_{\substack{t=-H \\ V(t)\neq 0}}^{H} \frac{1}{\sqrt{|V(t)|_P}} \tag{14}$$

zerlegbar. Es soll gezeigt werden, daß für jede dieser Summen eine Ungleichung der Form

$$\Sigma^* < C_{11} H^{n-1/2} \tag{15}$$

gilt und somit

$$\Sigma < C_{12} H^{2n-3/2} \tag{16}$$

ist.

Die Werte $|V(t)|$ sind ganze Zahlen x mit $1 \leqq x \leqq C_{13} H^{2n-2}$, da $D(q)$ ein homogenes Polynom vom Grad $2n-2$ in den Variablen $a_0, \ldots, a_n$ ist. Daher gilt, wenn die natürliche Zahl k durch die Ungleichung

$$P^k \leqq C_{13} H^{2n-2} < P^{k+1} \tag{17}$$

definiert wird, stets $|x|_P = P^{-\varkappa}$ mit $0 \leqq \varkappa \leqq k$, und zwar wird jedes solche $\varkappa$ für höchstens $A_\varkappa = C_{13} H^{2n-2} / P^\varkappa$ Werte x angenommen, da dann $P \mid x$ sein muß. Jede dieser Zahlen x, die auftritt, liefert zur Summe Σ^* den Beitrag $P^{\varkappa/2}$, und so ergibt sich durch Betrachtung des ungünstigsten Falles, wo die Zahlen x mit den größtmöglichen Werten von $P^{\varkappa/2}$ — und jeder von ihnen g mal — angenommen werden, die Abschätzung

$$\Sigma^* < g(A_k P^{k/2} + A_{k-1} P^{(k-1)/2} + \cdots + A_{k-u} P^{(k-u)/2}),$$

wenn u eine natürliche Zahl mit

$$A_k + A_{k-1} + \cdots + A_{k-u} \geqq 2H + 1 \tag{18}$$

ist. Daraus folgt weiter

$$\begin{aligned} \Sigma^* &< g C_{13} H^{2n-2} (P^{-k/2} + P^{-(k-1)/2} + \cdots + P^{-(k-u)/2}) \\ &= g C_{13} \frac{H^{2n-2}}{P^{k/2}} \cdot \frac{P^{u/2}-1}{P^{1/2}-1} < g C_{13}^{1/2} H^{n-1} P^{1/2} \cdot \frac{P^{u/2}-1}{P^{1/2}-1}, \end{aligned}$$

letzteres wegen (17). Folglich ist

$$\Sigma^* < C_{14} H^{n-1} P^{u/2}. \tag{19}$$

Ferner gilt auf Grund von (18)

$$2H+1 \leqq C_{13} H^{2n-2} (P^{-k} + P^{-k+1} + \cdots + P^{-u}) = C_{13} \frac{H^{2n-2}}{P^k} \cdot \frac{P^{u+1}-1}{P-1},$$

also, wieder nach (17),

$$P^{u+1} > C_{15} \frac{P^k H}{H^{2n-2}} > C_{15} C_{13} \frac{H}{P},$$

d.h. $P^u > C_{16} H$, und somit kann man $P^u = C_{17} H$ setzen, wenn $C_{17} > C_{16}$ geeignet gewählt wird. Damit ergibt sich aus (19) die Ungleichung (15), so daß (16) bewiesen ist.

Setzt man (16) in (12) ein, so erhält man

$$\overline{T}(S^3) \leqq C_{12} \varlimsup_{H_0 \to \infty} \sum_{H=H_0}^{\infty} H^{(2-\sigma)n-3/2},$$

und da in dieser Reihe wegen $\sigma = 2 - \frac{1}{2n} + \varepsilon$ der Exponent < -1 ist, tritt Konvergenz ein, so daß $\overline{T}(S^3) = T(S^3) = 0$ folgt. Damit ist alles bewiesen.

5. Beweis von Satz 2. Analog zum vorangehenden Beweis haben wir jetzt für $\sigma = \frac{3}{2} + \varepsilon$ mit beliebigem $\varepsilon > 0$ die Gleichung $T\big(S_2(\sigma)\big) = 0$ nachzuweisen. Die Darstellung von $S = S_2^*(\sigma)$ als Vereinigung von Mengen S^j vereinfacht sich jetzt zu $S = S^0 \cup S^1 \cup S^3$, wobei diese Mengen mit dem jetzigen Wert von σ in (8) definiert sind. Wegen $2\left(\frac{3}{2} + \varepsilon\right) > 2$ ergibt sich dann wie im Beweis von Satz 1 die Gleichung

$$T(S^0 \cap S^1) = 0.$$

Um die noch fehlende Aussage

$$T(S^3) = 0 \tag{20}$$

zu beweisen, können wir zunächst dem Beweisgang von Satz 1 bis zum Auftreten der Summe Σ in (13) folgen. Da jetzt für jedes Polynom $q(x) = a_0 x^2 + a_1 x + a_2$ offenbar $D(q) = a_1^2 - 4 a_0 a_2$ ist und mindestens einer der drei Koeffizienten den Betrag H haben muß, gilt die Ungleichung

$$\Sigma \leqq 2 \sum_{a_0, a_2 = -H}^{H}{}' |H^2 - 4 a_0 a_2|_P^{-1/2} + 4 \sum_{a_1, a = -H}^{H}{}' |a_1^2 - 4 H a|_P^{-1/2}. \tag{21}$$

Wir bezeichnen die beiden Summen auf der rechten Seite mit Σ_1 bzw. Σ_2. Zu ihrer Abschätzung berücksichtigen wir, daß die Werte von $D(q)$ ganze Zahlen x mit $1 \leqq |x| \leqq 5H^2$ sind und Gleichungen der Form $|x|_P = P^{-\varkappa}$ mit $0 \leqq \varkappa \leqq k$ genügen, wenn k wie in (17) (mit $C_{13} = 5$) definiert wird. Jedes solche $\varkappa$ wird somit höchstens $10H^2/P^\varkappa$ Werte von x angenommen, und ferner gibt Hilfssatz 9 eine obere Schranke für die Vielfachheit an, mit der eine Zahl x in der Summe Σ_1 auftreten kann. Somit erhält man

$$\begin{aligned}\Sigma_1 &\leqq 4H + \sum_{\substack{a_0, a_2 = -H \\ 4 a_0 a_2 \neq 0}}^{H}{}' |H^2 - 4 a_0 a_2|_P^{-1/2} \leqq 4H + C_2 H^{C_3/\log\log H}\, 10 H^2 \sum_{\varkappa=0}^{k} P^{-\varkappa/2} \\ &\leqq C_{18} H^{2 + C_{14}/\log\log H}.\end{aligned}$$

Ähnlich folgt auf Grund der Hilfssätze 10 und 11

$$\begin{aligned}\Sigma_2 &\leqq C_2 H^{C_3/\log\log H} \sum_{\varkappa=0}^{k} \sum_{x_1=1}^{[5H^2 P^{-\varkappa}]} r(x_1 P^\varkappa, 4H)\, P^{\varkappa/2} \\ &\leqq C_2 H^{C_3/\log\log H} \sum_{\varkappa=0}^{k} \sum_{x_1=1}^{[5H^2 P^{-\varkappa}]} r(x_1, 4H)\, P^\varkappa \leqq C_2 H^{C_3/\log\log H} (k+1)\, 5H^2 \log(4H) \\ &< C_{19} H^{2 + C_{20}/\log\log H}.\end{aligned}$$

Setzt man dies in (12) ein, so ergibt sich

$$\overline{T}(S^3) \geqq C_{21} \varliminf_{H_0 \to \infty} \sum_{H=H_0}^{\infty} H^{-2\sigma + 2 + C_{22}/\log\log H},$$

woraus die zu beweisende Gl. (20) folgt.

6. Schlußbemerkung. Im Reellen bzw. im Komplexen kann man bekanntlich die Mahlersche Vermutung so aussprechen[13]), daß für fast alle ξ die Gleichungen

$$\vartheta_n(\xi) = 1 \quad \text{bzw.} \quad \vartheta_n(\xi) = \frac{1}{2} - \frac{1}{2n} \qquad (n = 1, 2, \ldots)$$

gelten. Ersteres hat sowohl die Aussage

$$\vartheta(\xi) = \sup_{n=1,2,\ldots} \vartheta_n(\xi) = 1 \qquad \text{für fast alle reellen } \xi \tag{22}$$

als auch

$$w(\xi) = \overline{\lim_{n\to\infty}}\, \vartheta_n(\xi) = 1 \qquad \text{für fast alle reellen } \xi \tag{23}$$

zur Folge. Analog dazu wird man im P-adischen Fall vermuten dürfen, daß fast alle $\xi \in K_p$ den Gleichungen

$$\vartheta_n(\xi) = 1 + \frac{1}{n} \qquad (n = 1, 2, \ldots) \tag{24}$$

genügen, woraus einerseits

$$\vartheta(\xi) = 2 \qquad \text{für fast alle } \xi \in K_P \tag{25}$$

und andererseits

$$w(\xi) = 1 \qquad \text{für fast alle } \xi \in K_P \tag{26}$$

folgt. Für $n = 1$ und $n = 2$ ist (24) durch Satz 1 bzw. Satz 2 bewiesen, während für $n \geqq 3$ die entsprechende Frage noch offensteht. Die Richtigkeit der Aussage (25) dagegen ergibt sich als unmittelbare Folge von Satz 1. Jeder weitere Fortschritt in Richtung auf (26) bzw. (24) dürfte — wie im gewöhnlichen Fall — von neuen Einsichten in die Werteverteilung der Diskriminanten $D(q)$ abhängig sein.

Literatur

[1] FELLER, W., u. E. TORNIER: Maß- und Integrationstheorie des Baireschen Nullraumes. Math. Ann. **107**, 165—187 (1932). — [2] KASCH, F.: Über eine metrische Eigenschaft der S-Zahlen. Math. Z. **70**, 263—270 (1958). — [3] KASCH, F., u. B. VOLKMANN: Zur Mahlerschen Vermutung über S-Zahlen. Math. Ann. **136**, 442—453 (1958). — [4] KUBILYUS, J. F.: Über eine Anwendung der Winogradowschen Methode auf die Lösung eines Problems aus der metrischen Zahlentheorie. [Russ.] Dokl. Akad. Nauk SSSR, N.S. **67**, 783—786 (1949). — [5] MAHLER, K.: Über Diophantische Approximationen im Gebiete der P-adischen Zahlen. Jahresber. DMV **44**, 250—255 (1934). — [6] MAHLER, K.: Über eine Klasseneinteilung der P-adischen Zahlen. Mathematica (Zutphen) **3**, 177—185 (1934/35). — [7] SCHNEIDER, TH.: Einführung in die transzendenten Zahlen. Berlin-Göttingen-Heidelberg: Springer 1957. — [8] TURKSTRA, H.: Metrische bijdragen tot de theorie der Diophantische approximaties in het lichaam der P-adische getallen. Diss. Freie Univ. Amsterdam 1936, 146 S.

Heidelberg, Mathematisches Institut der Universität
Mainz, Mathematisches Institut der Universität

(Eingegangen am 29. April 1959)

[13]) Vgl. [3], § 1.

Moser, J.
Math. Zeitschr. 72, 379–398 (1960)

The Order of a Singularity in Fuchs' Theory

By

Jürgen Moser

I. Introduction

1. This paper is concerned with systems of ordinary differential equations

$$\frac{dy}{dx} = A(x)\, y \tag{1}$$

where y is a vector with n components and $A(x)$ an n by n matrix. The elements of $A(x)$ are assumed to be analytic functions of the complex variable x for $0<|x|<\varrho$ with a pole at $x=0$:

$$A(x) = x^{-p} \sum_{\nu=0}^{\infty} A_\nu x^\nu \quad (A_0 \neq 0) \tag{2}$$

where p is an integer and the A_ν constant matrices. The aim of Fuchs' theory is to study the nature of the solution vectors $y(x)$ near the singularity $x=0$.

In most treatments of this problem one restricts the attention to the case where A_0 is a matrix with distinct eigenvalues in which case one can describe the solution by asymptotic series for $p>1$. For $p=1$ one has a regular singularity, i.e. all solutions $y(x)$ of (1) grow at most like a finite power of $|x|$

$$|y| \leqq c|x|^N \tag{3}$$

for small x and with some norm $|y|$ [1].

For the degenerate case, however, where A_0 has several zero eigenvalues or even is nilpotent this problem is very complicated and has not been treated in this generality. It can happen, for instance, that even in the case $p>1$ one has a regular singularity (see example below). On the other hand it was proven by Horn [1] that in case $x=0$ is a regular singularity one can transform the system (1) by

$$y = T(x)\, z \tag{4}$$

into a system

$$\frac{dz}{dx} = B(x)\, z, \qquad B = T^{-1} A\, T - T^{-1} \frac{d}{dx} T \tag{5}$$

where $B(x)$ has a pole of first order only. Here $T(x)$ is a matrix whose elements are analytic in some neighborhood $0<|x|<\varrho$ with at most a pole at $x=0$, and which satisfies $\det T(x) \not\equiv 0$.

[1]) Since $y(x)$ is multivalued in general one has to restrict the argument of x to a fixed sector $|\arg x| < \text{const}$.

We will call systems (1) and (5) which are related by such a transformation (4) "equivalent", since the behavior of the solutions near a singularity is essentially the same, namely up to functions which have a pole near the origin. Here we require only that $\det T(x)$ is not identically zero but allow that $\det T(0)$ vanishes [2]).

Thus Horn's theorem can be expressed as follows: A system (1) has a regular singularity at $x=0$ if and only if it is equivalent to a system (2) with a pole of first order. Even though this criterion gives a necessary and sufficient condition for a regular singularity it cannot be used to decide about the nature of the singularity for a preassigned system. There is no method of constructing the transformation $T(x)$, and it is not even clear, how many terms of the transformation matrix

$$T(x) = \sum_{\nu=-N}^{M} x^\nu T_\nu$$

are necessary for the reduction. It is the purpose of this note to give a criterion which allows the decision of this question in a definite number of steps.

2. For any system of the form (1) we define the rational number

$$m(A) = p - 1 + \frac{r}{n} \quad \geqq 0 \tag{6}$$

as the *order* of A where $r = r(A_0)$ is the algebraic rank of A_0, $0 < r \leqq n$. If $p - 1 + \frac{r}{n} < 0$ we set $m(A) = 0$. For every system (1) we introduce the number

$$\mu(A) = \operatorname*{Min}_T m(T^{-1}AT - T^{-1}T'), \tag{6'}$$

i.e. $\mu(A)$ is the minimum value of the order of all systems which are equivalent to (1). One can consider $\mu(A) - 1$ as a generalization of "Poincaré's rank" of a singularity. The matrix $A(x)$ is called *reducible*, if $m(A) > \mu(A)$.

Before stating the main results we illustrate the significance of $\mu(A)$ for the nilpotent matrix

$$A(x) = \begin{pmatrix} x^{-q} & -x^{-2q} \\ 1 & -x^{-q} \end{pmatrix}$$

where q is a positive integer. Obviously $m(A) = 2q - \frac{1}{2}$ but this matrix is reducible. With the transformation matrix

$$T(x) = \begin{pmatrix} 1 & 0 \\ x^q & x^{q+\gamma} \end{pmatrix}, \quad \gamma = \left[\frac{q}{2}\right] \text{ [3])}$$

one computes $B = T^{-1}AT - T^{-1}T'$ to

$$B(x) = \begin{pmatrix} 0 & -x^{-q+\gamma} \\ -q\,x^{-\gamma-1} & -(q+\gamma)\,x^{-1} \end{pmatrix}.$$

[2]) A similar concept of equivalence was introduced by G. D. BIRKHOFF [2], who required that $T(0)$ is the identity matrix. In this case, A_0 is unchanged, and the order p of the pole is the same for equivalent systems.

[3]) $[x]$ stands for the greates integer $\leqq x$.

One can show that $B(x)$ is not reducible and one finds

$$\mu(A) = m(B) = \frac{q+1}{2}.$$

In particular the minimum order of the pole $x=0$ is

$$\gamma + 1 = \left[\frac{q}{2}\right] + 1.$$

In this simple example the order $m(A)$ can be reduced from $2q - \frac{1}{2}$ to $\frac{1}{2}(q+1)$.

The regular singularity is characterized by the inequality

$$\mu \leqq 1$$

according to Horn's theorem. In the following we will assume that

$$m(A) > 1$$

and investigate under which conditions $A(x)$ is reducible.

The reason for the above refined definition of the order $m(A)$ of the singularity lies in the fact that it is possible to give a necessary and sufficient criterion for reducibility which involves the first two coefficients A_0, A_1 of $A(x)$ only. This statement is contained in

THEOREM 1. *If $m(A)>1$ the system (1) is reducible if and only if the polynomial*

$$\mathfrak{P}(\lambda) = x^r \det\left(\lambda I + x^{p-1} A(x)\right)\big|_{x=0}, \qquad r = r(A_0) \tag{7}$$

vanishes identically in λ.

The polynomial $\mathfrak{P}(\lambda)$ depends on A_0, A_1 only, since in

$$x^{p-1} A(x) = \frac{A_0}{x} + A_1 + x A_2 + \cdots$$

A_0 has the rank r. Hence in forming the determinant in (7) one has to take r columns from A_0 and the others from $\lambda I + A_1 + x A_2 + \cdots|_{x=0} = \lambda I + A_1$. Thus

$$\mathfrak{P}(\lambda) = \sum_{\nu=0}^{n-r} \lambda^\nu \mathfrak{P}_\nu(A_0, A_1)$$

where the coefficients $\mathfrak{P}_\nu(A_0, A_1)$ are homogeneous polynomials in the coefficients of A_0, A_1 of degree $n-\nu$. Therefore the above criterion requires

$$\mathfrak{P}_\nu(A_0, A_1) = 0 \quad \text{for} \quad \nu = 0, \ldots, n-r.$$

Since reducibility depends on the coefficients A_0, A_1 one can expect that the reduction of $m(A)$ can be achieved by simple transformation matrices $T(x)$. This is expressed by

THEOREM 2. *If $A(x)$ is reducible and $m(A)>1$ then the reduction can be carried out with a matrix T of the form*

$$T=(P_0+xP_1)\operatorname{diag}(1,1,\ldots,1,x,x,\ldots x)$$

where P_0, P_1 are constant matrices and $\det P_0\neq 0$.

REMARK. This theorem expresses only that $B=T^{-1}AT-T^{-1}T'$ satisfies $m(B)<m(A)$, not that $m(B)$ equals $\mu(A)$.

Theorem 2 limits the computation of T to two matrices. After this reduction has been carried out one can apply Theorem 1 to check whether a further reduction is possible. Here it is important to observe that $\mathfrak{P}(\lambda)$ will change from one step to the next since B is not obtained by a similarity transformation but by $B=T^{-1}AT-T^{-1}T'$. If the system possesses a regular singularity one has to reduce the order from $m(A)$ to $\mu=1$ which requires at most $(p-2)n+r=n(m(A)-1)$ steps. In this sense one can consider Theorem 1 as a criterion for a regular singularity, although it is not given in "closed form". The same remark refers more generally to the computation of $\mu(A)$. After one has reached the last step for which $m(A)=\mu(A)$ one has $\mathfrak{P}(\lambda)\not\equiv 0$ which is the characteristic property of A_0, A_1 in the irreducible case.

3. It is interesting to note that for ordinary differential equation

$$\frac{d^nu}{dx^n}+a_1(x)\frac{d^{n-1}u}{dx^{n-1}}+\cdots+a_{n-1}(x)\frac{du}{dx}+a_n(x)\,u=0 \tag{8}$$

the invariant μ can be computed explicitly and not only recursively. Here the $a_k(x)$ are assumed to be analytic in $0<|x|<\varrho$ with at most a pole at $x=0$. The equation (8) can be rewritten as a first order system (1) for which the number $\mu=\mu(A)$ has been defined.

This number can be computed as follows: Let $\lambda_k\geqq 0$ be the smallest integer such that

$$x^{\lambda_k}a_k(x) \quad \text{is regular at} \quad x=0.$$

THEOREM 3. *With the above definition let p be the smallest integer with*

$$p\geqq\frac{\lambda_k}{k}, \qquad r=\operatorname*{Max}_k(\lambda_k-(p-1)k)>0.$$

Then

$$\mu=p-1+\frac{r}{n} \tag{9}$$

is the required invariant provided $p>1$. If $p\leqq 1$ one has $\mu\leqq 1$.

This theorem can be considered as a generalization of Fuchs' theorem which states that (8) has a regular singularity if and only if $\lambda_k\leqq k$, i.e. $p\leqq 1$. Such a criterion is not known for systems (1). A similar situation occurs in the computation of μ, which can be done explicitly for differential equations (8) while for systems only a "recursive" criterion can be given.

The differential equations (8) can be chosen as models to exhibit all non-negative rational numbers with denominator n as possible values for μ. For $0<\mu\leqq 1$ one has a regular singularity and for $\mu=0$ a regular point.

If the matrix $A(x)$ is real and symmetric or hermitean the computation of μ can be given explicitly since only the reduced case occurs.

4. The above question is closely related to the problem of invariants of $A(x)$ under the equivalence

$$A \sim B = T^{-1} A\, T - T^{-1} T' \tag{10}$$

where $T(x)$ is a matrix described under (4). One might expect similar normal forms as they are known under similarity transformations. That the situation is quite different in both cases is seen from the fact that the invariants under similarity consist of at least n power series (for instance the coefficients of the characteristic polynomial) while the equivalence (10) admits only finitely many numbers as invariants for A. This follows from a theorem of G. D. Birkhoff [2] who proved that one can find a $T(x)$ with $T(0)=I$ such that

$$B = \sum_{\nu=1}^{p} B_{p-\nu}\, x^{-\nu}$$

does not contain x in positive powers. His method requires solving an integral equation and might be called *transcendental*, since it requires the knowledge of infinitely many terms in A. In contradistinction to Birkhoff's result this paper is concerned with terms of negative exponents in x and an *algebraic* description of those terms. While in Birkhoff's paper the term $T^{-1}T'$ in (10) is to be considered as the principal part $T^{-1}A\,T$ is the main term in this note.

The equivalence relation (10) has been investigated in several papers, in particular, by Loewy [6], where the question of decomposability of a matrix or a matrix complex is treated. However, these results are not used and not needed for the proof of the above results.

In this paper we consider the coefficients of all analytic functions to be complex numbers although it is clear that the proofs can be generalized to commutative fields without zero divisors.

II. 3 Lemmata

The proof of the above theorem is based on 3 lemmata which are proven in this section. The first lemma can be considered as a generalization of the fact that an analytic function $f(z)\equiv 0$ which has at most a pole at $z=0$ can be written in the form $f(z)=z^{\alpha}g(z)$ where $g(z)$ is analytic at $z=0$ and $g(0)\neq 0$. The number α will be replaced by a diagonal matrices in the corresponding statement for matrices.

LEMMA 1. *Let*

$$T(x) = \sum_{\nu} T_\nu x^\nu$$

be a sum with finitely many negative ν *and* $\det T(x) \not\equiv 0$. *Then* $T(x)$ *can be represented in the form*

$$T(x) = P(x)\, x^\alpha\, Q(x)$$

where $P(x)$ *is a polynomial with* $\det P(x) \equiv 1$ *and* $Q(x)$ *a power series with* $\det Q(0) \neq 0$ *and*

$$x^\alpha = \operatorname{diag}(x^{\alpha_1}, \ldots, x^{\alpha_n})$$

where $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$ are integers.

REMARK 1. $P(x)$ is a unit in the ring of polynomials since on account of $\det P(x) \equiv 1$ the inverse $P^{-1}(x)$ also is a polynomial. $Q(x)$ is a unit in the ring of power series while x^α represents the essential part of $T(x)$.

REMARK 2. Like in Lemma 1 one can find a representation of the form

$$T(x) = Q(x)\, x^\alpha\, P(x)$$

with the same properties specified for P, Q in Lemma 1.

PROOF. We prefer to prove the statement in the form of remark 2

$$T(x) = Q(x)\, x^\alpha\, P(x)$$

considering the column vectors $t_l(x)$ of $T(x)$ instead of the rows. The matrix $P(x)$ will be built up by multiplication out of "elementary" matrices $E(x), F$:

$$(*) \qquad E = I + E_{kl}(x) \qquad (k \neq l)$$

where E_{kl} has all elements equal to zero but one element in the k-th row and l-th column which is any polynomial $p_{kl}(x)$, $k \neq l$. Obviously $\det E \equiv 1$. F stands for any permutation matrix

$$F = (\pm e_{j_1}, \pm e_{j_2}, \ldots, \pm e_{j_n})$$

where e_l is the l-th unit vector and $j_1, \ldots, j_n$ is a permutation of $1, 2, \ldots, n$. The signs $\pm$ are arbitrary except for the condition $\det F = 1$. Thus F is independent of x.

Multiplying a matrix

$$T = (t_1, t_2, \ldots, t_n)$$

with E from the right results in replacing t_l by $t_l + p_{kl}(x)\, t_k$ while multiplication with F from the right permuts the column vectors and changes the signs. Let P denote any product of several such matrices E and F. Obviously these matrices form a group all elements of which are matrices of polynomials with $\det P \equiv 1$.

To prove Lemma 1 we define the integers $\alpha_1, \alpha_2, \ldots, \alpha_n$ successively by the following maximal conditions: Let α_1 be the greatest integer such that $x^{-\alpha_1} T(x)$ is a power series. The constant term of $x^{-\alpha_1} T(x)$ then is a nonzero matrix and applying a matrix F we can assume that

$$T(x)\, F = \left(x^{\alpha_1} s_1(x), x^{\alpha_1} s_2(x), \ldots, x^{\alpha_1} s_n(x)\right)$$

where $s_l(x)$ are power series and $s_1(0) \neq 0$. Obviously α_1 remains unchanged if T is replaced by TP.

α_2 is chosen as the greatest integer under all choices of P such that the representation

$$T(x)\,P = \left(x^{\alpha_1} r_1(x),\ x^{\alpha_2} r_2(x),\ x^{\alpha_2} r_3(x), \ldots, x^{\alpha_2} r_n(x)\right)$$

holds where the $r_l(x)$ are power series. Obviously $\alpha_2 \geqq \alpha_1$. After having defined $\alpha_1, \ldots, \alpha_{l-1}$ we choose α_l to be the greatest integer for which a representation

$$T(x)\,P(x) = \left(x^{\alpha_1} q_1(x), \ldots, x^{\alpha_l} q_l(x),\ x^{\alpha_l} q_{l+1}(x), \ldots, x^{\alpha_l} q_n(x)\right) \tag{2.1}$$

holds. Again $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_l$ is an obvious consequence of the definition. The existence of α_l follows from the assumption that $\det\left(T(x)\,P(x)\right) = \det T(x) \not\equiv 0$, hence there is an integer γ with $T(x) = c\,x^{\gamma} + \cdots$, $c \neq 0$. The estimate $\alpha_1 + \alpha_2 + \cdots \alpha_{e-1} + (n - l + 1)\,\alpha_l \leqq \gamma$ guarantees the existence of a maximal α_l.

Now it follows by induction that in any maximal representation (2.1)

$$\operatorname{rank}\left(q_1(0), q_2(0), \ldots, q_n(0)\right) \geqq l. \tag{2.2}$$

For $l = 1$ this is obvious. Assume this has been proven for $l - 1$. If necessary one can achieve then that $q_1(0), q_2(0), \ldots, q_{l-1}(0)$ are linearly independent vectors by applying a permutation matrix F to (2.1). If the statement (2.2) were wrong then

$$q_k(0) = \sum_{\lambda=1}^{l-1} c_{k\lambda}\, q_\lambda(0) \quad \text{for} \quad k = l, \ldots, n.$$

Applying a product of matrices E we can achieve

$$q_k(0) = 0 \quad \text{for} \quad k = l, \ldots, n$$

which implies that α_l can be increased. This contradiction proves (2.2). For $l = n$ we obtain

$$T\,P = \left(x^{\alpha_1} q_1(x), \ldots, x^{\alpha_n} q_n(x)\right) = Q(x)\,x^{\alpha}$$

where

$$Q(x) = \left(q_1(x), \ldots, q_n(x)\right), \qquad x^{\alpha} = \operatorname{diag}\left(x^{\alpha_1}, \ldots, x^{\alpha_n}\right)$$

is a power series with

$$\det Q(0) \neq 0.$$

Thus

$$T = Q(x)\,x^{\alpha}\,P^{-1}$$

and since P^{-1} also is a polynomial with $\det P \equiv 1$ Lemma 1 is proven in the form of remark 2. Applying the argument to the transposed of T one obtains the statement of the Lemma 1.

The representation of the Lemma 1 is not unique, the numbers $\alpha_1, \ldots, \alpha_n$, however, depend on T only. For, assume that there is a second representation

$$T = \hat{P}\, x^{\beta}\, \hat{Q} = P\, x^{\alpha}\, Q.$$

Then, from the maximal choice of the α it follows

$$\beta_k \leqq \alpha_k.$$

and from

$$\det T = \det \hat{Q}\,(0)\; x^{\Sigma \beta_k} + \cdots = \det Q\,(0)\; x^{\Sigma \alpha_k} + \cdots$$

one has

$$\sum_{k=1}^{n} \beta_k = \sum_{k=1}^{n} \alpha_k.$$

These two relations imply $\beta_k = \alpha_k$.

One can give an independent description for the α_k: Consider any subdeterminant Δ of order l of T and let

$$\Delta = c\, x^{\gamma_l} + \cdots \qquad c \neq 0$$

be the expansion of Δ. Then

$$\alpha_1 + \cdots + \alpha_l = \operatorname{Min} \gamma_l$$

where the minimum is taken over all l by l subdeterminants of T. This defines the α_k. The proof of this statement follows immediately from Lemma 1.

For any matrix $T(x)$ of the type occurring we define the span $\sigma(T)$ by

$$\sigma(T) = \alpha_n - \alpha_1 = \operatorname*{Max}_{k,\,l} (\alpha_k - \alpha_l)$$

which is a non negative integer. The assumption $\sigma(T) = 0$, for example, amounts to

$$T(x) = x^{\alpha_1}\, Q(x); \qquad \det Q\,(0) \neq 0.$$

Any matrix T of span $\sigma > 0$ can be written as the product of σ matrices of span 1. For this purpose observe that x^{α} can be broken up into σ matrices of span 1 by constructing σ sequences $\beta^{(\nu)} = (\beta_k^{(\nu)})$ $(k = 1, \ldots, n;\ \nu = 1, \ldots, \sigma)$ with

$$\beta_k^{(\nu)} \leqq \beta_{k+1}^{(\nu)}; \qquad \beta_n^{(\nu)} - \beta_1^{(\nu)} = 1.$$

and

$$\sum_{\nu=1}^{\sigma} \beta_k^{(\nu)} = \alpha_k.$$

Defining

$$T_\nu = P\, x^{\beta^{(\nu)}}\, P^{-1} \quad \text{for} \quad \nu = 1, 2, \ldots, \sigma - 1$$

$$T_\sigma = P\, x^{\beta^{(\sigma)}}\, Q$$

one has

$$\sigma(T^{(\nu)}) = 1.$$

$$T = T_1\, T_2 \ldots T_\sigma.$$

For later reference we introduce the ring $R[x]$ of functions $f(x)$ which are analytic in some (individual) neighborhood $|x|<\varrho$. $R(x)$ denotes the corresponding field, generated by $R[x]$. The elements in $R(x)$ can be written in the form $x^\alpha f(x)$ where α is an integer and $f\in R[x]$, $f(0)\neq 0$. With $M[x]$, $M(x)$ we denote the ring of matrices whose elements lie in $R[x]$, $R(x)$ respectively. Thus $M[x]$, $M(x)$ formnon commutative rings with zero divisors. The unit elements in the ring $M[x]$, consist of those matrices $Q(x)$ for which $\det Q(0)\neq 0$.

Similarly let $R_0[x]$ denote the ring of polynomials of x and $R_0(x)$ the field generated by $R_0[x]$. If $M_0[x]$, $M_0(x)$ are the rings of matrices whose elements are in $R_0[x]$, $R_0(x)$ one can express Lemma 1 as follows: Every $T(x)\in M(x)$ can be written in the form $P(x)\,x^\alpha\,Q(x)$ where P is a unit in $M_0[x]$ and Q is a unit in $M[x]$. x^α has the same meaning as above.

Lemma 2. *Let*

$$A(x)=A_0+x\,A_1+\cdots$$

belong to $M[x]$ *and satisfy* $r=r(A_0)>0$ (i.e. $A_0\neq 0$). *Let* $T(x)$ *be a matrix in* $M(x)$ *with* $\det T(x)\not\equiv 0$ *and define*

$$B(x)=T^{-1}A\,T.$$

A necessary and sufficient condition for the existence of such a $T(x)$ *that*

$$B(x)=B_0+x\,B_1+\cdots$$

belongs to $M[x]$ *and satisfies*

$$r(B_0)<r(A_0)$$

is that the polynomial

$$\mathfrak{P}(\lambda)=x^r\det\left(\lambda\,I+\frac{A(x)}{x}\right)\Big|_{x=0}$$

vanishes identically in λ.

Proof. This condition is certainly necessary: If there exists such a $T(x)$ then

$$\det\left(\lambda\,I+\frac{A(x)}{x}\right)=\det\left(\lambda\,I+\frac{B(x)}{x}\right)=$$

$$=\det\left(\lambda\,I+\frac{B_0}{x}+B_1+\cdots\right).$$

The last determinant contains x^{-1} at most to the power $r(B_0)<r$, hence

$$x^r\det\left(\lambda\,I+\frac{A(x)}{x}\right)$$

vanishes identically in λ for $x=0$.

To prove the converse we make use of the theory of invariants of matrices under similarity transformations [4]).

[4]) The following proof is a slight modification of an argument for which I am indebted to Professor A. A. Albert.

Let $$f(\lambda, x) = \det\left(\lambda I + A(x)\right)$$

be the characteristic polynomial of $A(x)$. Then the condition $\mathfrak{P}(\lambda) \equiv 0$ is tantamount to

$$x^{n-r+1} | f(\lambda x, x)$$

i.e. to the assumption that $x^{-n+r-1} f(\lambda x, x)$ is a power series in x. This follows immediately from the equation

$$x^{-n+r-1} f(\lambda x, x) = x^{r-1} \det\left(\lambda I + \frac{A(x)}{x}\right).$$

Denote by s the largest integer for which $x^s | f(\lambda x, x)$. Then $n - r < s \leqq n$. We will prove the converse of Lemma 2 in the stronger version, that the rank of B_0 satisfies

(2.2*) $$r(B_0) \leqq n - s;$$

which implies $r(B_0) < r(A_0)$.

To prove this statement we make use of the fact that a matrix $C = C(x)$ in $M(x)$ is similar to $A(x)$ if and only if it has the same elementary divisors $f_j(\lambda, x)$ as $A(x)$. The $f_j(\lambda, x)$ are polynomials in λ with coefficients in the field $R(x)$ which satisfy

(2.3) $$f(\lambda, x) = \prod_{j=1}^{\omega} f_j(\lambda, x).$$

Since the highest coefficient of $f_j(\lambda, x)$ is one and concludes from Gauss' Lemma that the coefficients of $f_i(\lambda, x)$ lie in the ring $R[x]$.

Corresponding to the factorization of $f(\lambda, x)$ one can decompose the n dimensional vector space into n_j dimensional subspaces $\left(n = \sum_{j=1}^{\omega} n_j\right)$ which are invariant under A. In the invariant subspaces A is given by matrices A_j with respect to an appropriate basis and

$$f_j(\lambda, x) = \det\left(\lambda I_{n_j} + A_j\right).$$

Let s_j be the largest integer satisfying $x^{s_j} | f_j(\lambda x, x)$ then it follows from (2.3) that $\sum_{j=1}^{\omega} s_j = s$.

Assuming that the statement (2.2*) has been proven for the indecomposable matrices A_j[5]), namely that they are similar to $B_j(x)$ in $M[x]$ with $r\left(B_j(0)\right) \leqq n_j - s_j$ it follows that the matrix $B(x)$, which is composed of the $B_j(x)$, satisfies

$$r\left(B(0)\right) = \sum_{j=1}^{\omega} r\left(B_j(0)\right) \leqq \sum_{j=1}^{\omega} (n_j - s_j) = n - s.$$

Therefore it suffices to prove the statement (2.2*) for indecomposable matrices $A(x)$. Then

(2.4) $$f(\lambda, x) = \lambda^n + a_n(x)\,\lambda^{n-1} + \cdots + a_1(x)$$

[5]) For the concepts used here compare [5], Chap. IV.

is the minimal polynomial and the condition $x^s/f(\lambda x, x)$ ensures that

$$b_1(x) = \frac{a_1}{x^s}, \quad b_2 = \frac{a_2}{x^{s-1}}, \ldots, b_s = \frac{a_s}{x}, \quad b_{s+1} = a_{s+1}, \ldots, a_n = b_n.$$

are power series in $R[x]$. Any indecomposable matrix $A(x)$ satisfying (2.4) is similar to the so-called companion matrix $C(x)$ of $f(\lambda, x)$ which contains ones in the diagonal above the main diagonal and has as last row the elements

$$(a_1, a_2, \ldots, a_n) = (b_1 x^s, b_2 x^{s-1}, \ldots, b_s x, b_{s+1}, \ldots, b_n).$$

All other elements of C are 0.

Defining the diagonal matrix

$$D(x) = (x^{-s}, x^{-s+1}, \ldots, x^{-1}, 1, \ldots, 1)$$

and

$$D^{-1} C D = B(x)$$

it is easily seen that $B(x)$ belongs to $M[x]$ and that the first s rows of $B(0)$ are zero rows, hence

$$r\big(B(0)\big) \leqq n - s$$

which proves the statement (2.2*) and the Lemma 2.

LEMMA 3. *Let $A(x)$ satisfy the assumption of Lemma 2 and assume*

$$\mathfrak{P}(\lambda) \equiv 0.$$

Then there exists a transformation $T(x)$ (as in Lemma 2) for which

$$T^{-1} A T = B(x) \varepsilon M[x] \quad r\big(B(0)\big) < r\big(A(0)\big)$$

for which

$$\sigma(T) = 1.$$

PROOF. a) For the proof it suffices to assume that $A(0)$ is nilpotent. Otherwise $A(0)$ has a nonzero eigenvalue $a \neq 0$ and with a constant similarity one can achieve that $A(0)$ is a triangular matrix and $a_{11} = a \neq 0$. Let $\hat{A}$ be the $(n-1)$ by $(n-1)$ matrix which is obtained from $A(x)$ by cancelling the first row and column. It is easily seen that

$$\mathfrak{P}(\lambda) = x^r \det\left(\lambda I + \frac{A(x)}{x}\Big|_{x=0}\right) = a\,\hat{\mathfrak{P}}(\lambda)$$

where

$$\hat{\mathfrak{P}}(\lambda) = x^{r-1} \det\left(\lambda I + \frac{\hat{A}(x)}{x}\Big|_{x=0}\right.$$

and thus $\hat{A}$ satisfies the same hypothes as A does. If the statement can be proven for all matrices with $< n$ rows and columns it follows for n by n matrices provided there is a nonzero eigenvalue of $A(0)$. This reduces the proof to nilpotent $A(0)$.

b) By Lemma 2 there exists some $T(x) \varepsilon M(x)$ which diminishes the rank of the constant term. Representing $T(x)$ in the form

$$T(x) = P(x)\, x^{\alpha}\, Q(x)$$

according to Lemma 1 we define

$$\hat{A} = P^{-1} A\, P, \qquad \hat{B} = Q\, B\, Q^{-1}$$

so that

$$x^{-\alpha}\, \hat{A}\, x^{\alpha} = \hat{B}.$$

Since

$$r\left(\hat{A}(0)\right) = r\left(A(0)\right); \qquad r\left(\hat{B}(0)\right) = r\left(B(0)\right)$$

we can assume $T(x) = x^{\alpha}$, where

$$\alpha = \operatorname{diag}(\alpha_1, \ldots, \alpha_n), \qquad \alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n.$$

Furthermore, since $\mathfrak{P}(\lambda)$ depends on A_0, A_1 only we can assume $A_2 = A_3 = \cdots = 0$.

c) In the decomposition $P\, x^{\alpha} Q$ the factors P, Q are not uniquely determined since there are polynomials P_1, P_2 for which

$$P_1\, x^{\alpha} = x^{\alpha} P_2$$

for instance, for all P_1 which are upper triangular matrices $x^{-\alpha} P_1 x^{\alpha}$ are polynomials in x again.

This freedom in P, Q can be used to achieve that the nonzero column ctors of $A(0)$ are linearly independent. Thus if $A(0)$ is decomposed into several matrices with n rows:

$$A(0) = (\tilde{A}_1, \ldots, \tilde{A}_{\tau})$$

the rank is additive

$$r\left(A(0)\right) = \sum_{\nu=1}^{\tau} r(\tilde{A}_{\nu}).$$

To achieve this aim, we write

$$A(0) = (a_{kl})$$

and notice that

$$a_{kl} = 0 \quad \text{for} \quad \alpha_k > \alpha_l,$$

since otherwise $x^{-\alpha} A\, x^{\alpha}$ would have a singularity at 0. We represent $A(0)$ by blocks of matrices by combining the a_{kl} with $\alpha_k = \alpha$, $\alpha_l = \alpha'$ to $A_{\alpha\alpha'}$ so that

$$A(0) = (A_{\alpha\alpha'})$$

where α, α' runs over certain integers. The above condition amounts to

$$A_{\alpha\alpha'} = 0 \quad \text{for} \quad \alpha > \alpha'.$$

The diagonal matrices $A_{\alpha\alpha}$, of course, are nilpotent since $A(0)$ is. By a constant transformation

$$T_0 = (T_{0\alpha\alpha'})$$

which is a diagonal matrix in this block representation ($T_{\alpha\alpha'}=0$ for $\alpha \neq \alpha'$) one can achieve that $A_{\alpha\alpha}$ are in upper triangular normal form. Thus $A(0)=(a_{kl})$ are upper triangular with zero diagonal elements.

Now we achieve that the column vectors of $A(0)$ are replaced by 0 vectors and r independent vectors. For this purpose let a_l denote the column vectors of $A(0)$;

$$A(0)=(a_1, a_2, \ldots, a_n).$$

Let $a_1, a_2, \ldots, a_{l-1}$ consist of zero and independent vectors already and assume a_l is dependent on $a_1, a_2, \ldots, a_{l-1}$

$$a_l=\sum_{\lambda<l} c_\lambda a_\lambda.$$

Let C denote the constant matrix which is obtained from the unit matrix by replacing the l^{th} column by

$$\begin{pmatrix} -c_1 \\ -c_2 \\ \vdots \\ -c_{l-1} \\ 1 \\ 0 \\ \cdot \\ 0 \end{pmatrix}.$$

Then

$$A(0)\,C=(a_1, a_2, \ldots, a_{l-1}, 0, a_{l+1}, \ldots, a_n).$$

Since $C^{-1}A(0)C$ is obtained from $A(0)C$ by adding the l^{th} row to the λ^{th} ($\lambda<l$) the first l columns are not affected by this procedure since $A(0)C$ like $A(0)$ is upper triangular with zero diagonal elements. Also since

$$x^{-\alpha}C\,x^{\alpha}=C_1(x)$$

is a polynomial with $\det C_1(x)\equiv 1$ (hence a unit in the polynomial ring) we have

$$x^{-\alpha}(C^{-1}A\,C)\,x^{\alpha}=C_1^{-1}B\,C_1=\hat{B}$$
$$r\big(C^{-1}A(0)\,C\big)=r\big(A(0)\big)$$
$$r\big(\hat{B}(0)\big)=r\big(B(0)\big).$$

Thus we can assume that A is replaced by $C^{-1}A\,C$ in the above statement.

Thus we can assume that the rank is additive with respect to a decomposition of $A(0)$ into several column matrices.

d) For the sake of simplicity we first carry out the proof of Lemma 3 in case

$$\alpha_1<\alpha_2<\cdots<\alpha_n.$$

The general case will be a generalization of this idea.

By a), b), c) we can assume

$$A = A_0 + xA_1$$

A_0 consists of 0 and independent column vectors

A_0 is upper triangular.

Since

$$B = x^{-\alpha} A\, x^{\alpha}$$

$$B_{kl} = x^{-\alpha_k+\alpha_l} A_{kl}(x)$$

the elements of $B(0)$ are zero except if $-\alpha_k+\alpha_l=-1$. Thus with

$$B(0) = (b_{kl})$$

one has

$$b_{kl} = 0 \quad \text{for} \quad \alpha_k \neq \alpha_l + 1$$

in particular,

$$b_{kl} = 0 \quad \text{for} \quad k \neq l + 1.$$

The α_k form an increasing sequence of integers. We construct all those sequences $\beta_k^{(\nu)}$ for which $\beta_n^{(\nu)} - \beta_1^{(\nu)} = 1$ which have the jump from $k=\nu-1$ to $k=\nu$ $(\nu=2, \ldots, n)$. Introducing

$$\beta^{(\nu)} = \operatorname{diag}(\beta_1^{(\nu)}, \beta_2^{(\nu)}, \ldots, \beta_n^{(\nu)}) \qquad (\nu = 2, \ldots, n)$$

we form the matrices

$$x^{-\beta^{(\nu)}} A(x)\, x^{\beta^{(\nu)}} = C^{(\nu)}(x)$$

and claim that

(*)
$$r\left(C^{(\nu)}(0)\right) < r(A_0))$$

for some ν. Counterexamples show that in general this inequality is not true for all ν.

To prove this statement notice that $\beta_k^{(\nu)} - \beta_l^{(\nu)} > 0$ implies $\alpha_k - \alpha_l > 0$. Therefore $C^{(\nu)}(x)$ has no negative powers of x and the matrix $C^{(\nu)}(0)$ has the form

$$\begin{array}{c|c}
\begin{matrix} 0 & a_{12} \cdots a_{1,\nu-1} \\ \vdots & \ddots \quad \vdots \\ & a_{\nu-2,\nu-1} \\ 0 & \cdots \quad 0 \end{matrix} & \begin{matrix} \\ \\ \\ 0 \end{matrix} \\
\hline
\begin{matrix} & b_{\nu,\nu-1} \end{matrix} & \begin{matrix} 0 & a_{\nu,\nu+1} \\ & \ddots \\ & & a_{n-1,n} \\ & & 0 \end{matrix}
\end{array}$$

Hence the rank of $C^{(\nu)}(0)$ is at most

$$r(a_1, a_2, \ldots a_{\nu-1}, a_{\nu+1}, \ldots, a_n) + r(b_\nu')$$

where b_ν' stands for the ν^{th} row of $B(0)$. Since the rank is additive for A_0 it follows

$$r(C^{(\nu)}) \leqq r(A_0) - r(a_\nu) + r(b_\nu).$$

Adding over all ν one obtains

$$\sum_{\nu=2}^{n} r(C^{(\nu)}) \leqq (n-1)\, r(A_0) - r(A_0) + r\big(B(0)\big).$$

The assumption $r(A_0) > r\big(B(0)\big)$ leads to

$$\frac{1}{n-1} \sum_{\nu=2}^{n} r(C^{(\nu)}) < r(A_0)$$

which proves the statement (*).

e) A similar argument as in d) can be applied in the general case

$$\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n.$$

In this case, however, one first has to achieve that $B(0)$ has the additive rank property with respect to the rows and simultaneously $A(0)$ with respect to the columns. Here we refer to rows and columns in the block representation determined by indices k with equal α_k. In other words: Let

$$A_0 = (A^{(1)}, A^{(2)}, \ldots, A^{(N)})$$

be a decomposition of A_0 into N matrices of consecutive columns in such a manner that $\alpha_k \neq \alpha_{k'}$ if k, k' are the column numbers belonging to different matrices $A^{(K)}$. Then

$$r(A_0) = \sum_{K=1}^{N} r(A^{(K)}).$$

A similar property can be achieved simultaneously with respect to the *rows* of $B(0)$.

To prove this statement we can assume $A(0)$ to be an upper triangular matrix consisting of r independent column vectors and $n-r$ zero columns. Writing $A(0)$ in the block representation

$$A(0) = (A_{\alpha\alpha'}), \qquad B(0) = (B_{\alpha\alpha'})$$

where α, α' runs over the integers which are assumed by $\alpha_1, \ldots, \alpha_n$ one has in particular, $A_{\alpha\alpha'} = 0$ for $\alpha > \alpha'$. Breaking up $A(0)$ into column matrices

$$A_{\alpha'} = (A_{\alpha\alpha'})$$

one obviously has

(**) $$r\big(A(0)\big) = \sum_{\alpha'} r(A_{\alpha'}).$$

In order to achieve the corresponding property with respect to the rows of $B(0)$ we apply a constant matrix

$$C = (C_{\alpha\alpha'}) \qquad (C_{\alpha\alpha'} = 0 \quad \text{if} \quad \alpha \neq \alpha')$$

(diagonal in block representation) so that $A_{\alpha\alpha} = B_{\alpha\alpha}$ and hence $B(0)$ become lower triangular matrices. Since C commutes with x^α one has

$$x^{-\alpha}(C^{-1}AC)\,x^\alpha = C^{-1}BC$$

and the matrix $C^{-1}A(0)C$ satisfies (**) since the rank of $A_{\alpha'}$ remains unchanged.

Applying the construction of c) we achieve that $B(0)$ consists of zero and linear independent row vectors. For this purpose we construct a constant matrix C_1 such that

$$C_1 B(0) C_1^{-1}$$

consists of zero rows and linearly independent rows. C_1 has ones in the diagonal and zeros above the diagonal. Then $x^{-\alpha}Ax^\alpha = B$ goes over into

$$x^{-\alpha} C_2 A C_2^{-1} x^\alpha = C_1 B C_1^{-1}$$

where

$$C_2(x) = x^\alpha C_1 x^{-\alpha}$$

$C_2(x)$ is regular at $x=0$, since C_1 has zeros above the diagonal. Moreover $C_2(0)$ is nonsingular and $\det C_2(0) = 1$. Actually $C_2(0)$ is diagonal in the block representation so that the condition (**) will be satisfied for $C_2 A C_2^{-1}$ as for A. This shows that one can assume that A as well as B have the additive rank property with respect to their columns (A) and their rows (B) respectively.

f) Now the proof can be finished like in d). Let $\beta^{(\nu)} = \operatorname{diag}(\beta_1^{(\nu)}, \beta_2^{(\nu)}, \ldots, \beta_n^{(\nu)})$ $(\nu = 2, \ldots, N)$, where $\beta_k^{(\nu)}$ are increasing sequences of 0 and 1 for which the jump occurs only at a jump of the given sequence α_k. Then $T = x^{\beta^{(\nu)}}$ have span 1. Introducing

$$x^{-\beta^{(\nu)}} A(x)\, x^{\beta^{(\nu)}} = C^{(\nu)}(x) \qquad (\nu = 2, \ldots, N)$$

it can be shown that

$$r\left(C^{(\nu)}(0)\right) = r(A_0) - r(A^{(\nu)}) + r(B^{(\nu)})$$

where

$$A_0 = (A^{(1)}, A^{(2)}, \ldots, A^{(N)}), \qquad B_0 = \begin{pmatrix} B^{(1)} \\ \vdots \\ B^{(N)} \end{pmatrix}.$$

Adding these relations one finds

$$\frac{1}{N-1} \sum_{\nu=2}^{N} r\left(C^{(\nu)}(0)\right) < r(A_0)$$

hence

$$r\left(C^{(\nu)}(0)\right) < r(A_0)$$

for at least one ν. Since $x^{\beta^{(\nu)}}$ has span 1 the Lemma 3 is proven.

III. Proof of Theorem 1 and 2

a) Assume $A = x^{-p}(A_0 + x A_1 + \cdots)$ satisfies

$$\mathfrak{P}(\lambda) = x^r \det\left(\frac{A_0}{x} + A_1 + \lambda I\right)\Big|_{x=0} \equiv 0.$$

Then according to Lemma 3 there is a transformation

$$T(x) = T_0 + x T_1 + \cdots = T$$

with $\sigma(T) = 1$ such that

$$T^{-1} A T = B$$

and

$$m(A) > m(B).$$

Since $\sigma(T) = 1$ one can write T in the form $P x^\alpha Q$ with $\alpha_n - \alpha_1 = 1$. Hence

$$T^{-1} \frac{d}{dx} T = Q^{-1} x^{-\alpha} P^{-1} \frac{dP}{dx} x^\alpha Q + Q^{-1} \frac{\alpha}{x} Q + Q^{-1} \frac{dQ}{dx}$$

has at most a pole of order 1. Thus, if $p > 1$ (i.e. $m > 1$) then

$$m\left(T^{-1} A T - T^{-1} \frac{d}{dx} T\right) < m(A)$$

which proves the first part of the theorem.

b) If on the other hand there exists some $T(x)$ such that

$$T^{-1} A T - T^{-1} \frac{d}{dx} T = B$$

and

$$m(B) < m(A)$$

then one can represent T in the form

$$P x^\alpha Q.$$

Then with

$$P^{-1} A P - P^{-1} \frac{d}{dx} P = \hat{A}$$

$$Q B Q^{-1} + \left(\frac{d}{dx} Q\right) Q^{-1} = \hat{B}$$

one has

$$x^{-\alpha} \hat{A} x^\alpha - \frac{\alpha}{x} = \hat{B}$$

and

$$m(\hat{A}) = m(A); \qquad m(\hat{B}) = m(B)$$

since P, Q are regular at $x = 0$. Since we assumed $m(A) > 1$ it follows

$$m(\hat{A}) > m\left(\hat{B} + \frac{\alpha}{x}\right)$$

thus m can be diminished by a similarity transformation. This proves that

$$x^r \det(\lambda I + x^{(p-1)} \hat{A})|_{x=0} \equiv 0.$$

Since this expression depends on A_0, A_1 only and since $\hat{A} = -P^{-1}AP = -P^{-1}\frac{d}{dx}P$ is a power series in x it follows that $\mathfrak{P}(\lambda) = x^r \det(\lambda I + x^{p-1}A)|_{x=0} \equiv 0$. In other words: If the number $m(A)$ is greater than 1 it can be diminished by similarity if and only if it can be diminished by the operation

$$T^{-1}AT - T^{-1}T'.$$

The transformation T can always be chosen as to be of span 1. Moreover, representing T in the form $Px^\alpha Q$ of Lemma 1 one observes that $Q(x)$ does not change $m(B)$ and therefore can be replaced by the identity. Finally it is obvious that only the first two terms of $P(x)$ enter in B_0, so that we can assume $T = (P_0 + P_1 x)x^\alpha$. This completes the proof of theorem 2.

In particular one notices that for symmetric matrices $A(x)$ only the reduced form can occur. Since for a symmetric matrix the rank agrees with the number of non zero eigenvalues one has

$$\mathfrak{P}(\lambda) = x^r \det\left(\lambda I + \frac{A_0}{x} + A_1\right)\Big|_{x=0} = a_1, \ldots, a_r \det(\lambda I + \hat{A})$$

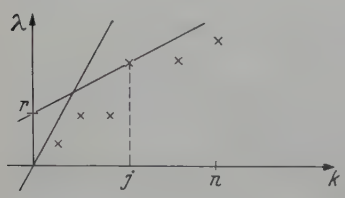
if $a_1, \ldots, a_r$ are the non zero eigenvalues of A_0. This expression is a polynomial of degree $n - r$.

IV. The case of a single differential equation

Consider a differential equation of n-th order

$$\text{(E)} \qquad \frac{d^n u}{dx^n} + a_1(x)\frac{d^{n-1}u}{dx^{n-1}} + \cdots + a_{n-1}(x)\frac{du}{dx} + a_n(x)u = 0$$

for which the coefficients $a_k(x)$ have at most a pole at zero. Let $\lambda_k \geqq 0$ denote the order of the pole of a_k. ($\lambda_k = 0$ if a_k is regular).



We now construct p, r as the smallest integers (p, r) ordered lexicographically such that

$$\lambda_k \leqq k(p-1) + r.$$

This can be visualized graphically if one plots the λ_k as ordinates over the abscissa k and then first constructs the straight line through 0 with the smallest integer slope p which stays above all λ_k. Then one finds among all the parallel lines of slope $p-1$ the lowest one. The section which is cut off on the λ-axis is r, which is an integer. That r is an integer follows from the formula

$$r = \operatorname{Max}_k(\lambda_k - (p-1)k).$$

Defining

$$\gamma_k = \operatorname{Min}(pk, (p-1)k + r)$$

we have

$$\lambda_k \leqq \gamma_k$$

and equality for at least one $k \geqq r$. Let $j \geqq r$ be the smallest integer with

$$\lambda_j = \gamma_j .$$

Geometrically γ_k represents the broken line dominating λ_k.

Now it is easy to construct an irreducible system of differential equations which is equivalent to (E). For this purpose introduce the vector y with the components

$$y_{k+1} = x^{-\gamma_{n-k}} \frac{d^k u}{dx^k} \qquad (k = 0, \ldots, n-1).$$

Then for $k = 0, \ldots, n-2$

$$\frac{dy_{k+1}}{dx} = x^{-\gamma_{n-k}+\gamma_{n-k-1}} y_{k+2} - \frac{\gamma_{n-k}}{x} y_{k+1}$$

$$\frac{dy_n}{dx} = -\sum_{k=0}^{n-1} a_{n-k} x^{-\gamma_1+\gamma_{n-k}} y_{k+1} - \frac{\gamma_1}{x} y_n .$$

Therefore, introducing

$$\gamma = \operatorname{diag}(\gamma_n, \ldots, \gamma_1)$$

and

$$K(x) = \left(\begin{array}{c|cc} \overbrace{\begin{array}{cccc} 0 & x & & \\ & \ddots & & \\ & & \ddots & \\ & & & x \\ & & 0 & \end{array}}^{n-r+1} & & \\ & \begin{array}{c} 1 \\ 0 \end{array} & \\ & & \begin{array}{cc} & 1 \\ b_{r-1} & b_1 \end{array} \\ b_n,\ b_j \ldots b_r & & \end{array}\right)$$

where the b_k are the regular functions

$$b_k(x) = -a_k(x)\, x^{\gamma_k} \qquad (k = 1, \ldots, n)$$

we have

$$\frac{dy}{dx} = \left(x^{-p} K(x) - \frac{\gamma}{x}\right) y \text{ [6]}.$$

It has to be shown that for

$$A(x) = x^{-p} K(x) - \frac{\gamma}{x}$$

the polynomial $\mathfrak{P}(\lambda)$ does not vanish identically, if $p > 1$. It is obvious that

$$x^p A(x)\big|_{x=0} = A_0 = K(0)$$

has rank r, since $r-1$ rows of $K(0)$ contains ones and since

$$b_j(x) = -a_j(x)\, x^{\gamma_j} = -a_j(x)\, x^{\lambda_j} \qquad (j \geqq r)$$

---

[6]) Use $\gamma_1 = \operatorname{Min}\{p, p-1+r\} = p$.

does not vanish at $x=0$. If one considers the determinant

$$\mathfrak{P}(\lambda) = x^r \det\left(\lambda I + \frac{1}{x} K(x) - \gamma x^{p-2}\right)\Big|_{x=0}$$

one computes easily the highest coefficient to be

$$\mathfrak{P}(\lambda) = \lambda^{n-j} b_j(0) + \cdots$$

where $j \geqq r$ was defined above. Hence the degree of $\mathfrak{P}(\lambda)$ is $\leqq n-r$ and $\mathfrak{P}(\lambda) \not\equiv 0$. This proves that the above system is not reducible, hence μ is given by

$$\mu = p - 1 + \frac{r}{n} \quad \text{if} \quad p > 1$$

where p, r were defined by the above diagram.

Necessary and sufficient for a regular singularity is that $p=1$ or $\mu \leqq 1$ which is the principal result of Fuchs theory. The above result can be considered a generalization of Fuchs' theorem.

References

[1] HORN, J.: Zur Theorie der Systeme linearer Differentialgleichungen mit einer unablängigen Veränderlichen, II. Math. Annalen **40**, 527–550 (1892). — [2] BIRKHOFF, G. D. : Equivalent singularities of ordinary linear differential equations. Math. Ann. **74**, 134–139 (1913). — [3] BIEBERBACH, L.: Theorie der gewöhnlichen Differentialgleichungen. Berlin: Springer 1953; esp. § 6. — [4] CODDINGTON, E. A., and N. LEVINSON: Theory of Ordinary Differential Equations. McGraw-Hill 1955; Chap. 4 and 5. — [5] ALBERT, A. A.: Modern Algebra. Chicago 1937, 1958; in particular Chap. IV. — [6] LOEWY, A.: Über Matrizen- und Differentialkomplexe, I, II, III. Math. Ann. **78**, 1–51, 343–358, 359–369 (1918). — [7] PERRON, O.: Über diejenigen Integrale linearer Differentialgleichungen, welche sich an einer Unbestimmtheitsstelle bestimmt verhalten. Math. Ann. **70**, 1–32 (1911).

Dept. of Math., Massachusetts Institute of Technology, Cambridge 39, Mass. (U.S.A.)

(Eingegangen am 2. Mai 1959)

Einführung in die Physik

Von R. W. Pohl, em. Professor der Physik an der Universität Göttingen

Erster Band: **Mechanik, Akustik und Wärmelehre**

14., verbesserte und ergänzte Auflage. Mit 575 Abbildungen, darunter 15 entlehnten. XII, 345 Seiten Gr.-8°. 1959. Ganzleinen DM 33,—

INHALTSÜBERSICHT

Mechanik: Einführung, Längen- und Zeitmessung. Darstellung von Bewegungen, Kinematik. Grundlagen der Dynamik. Anwendung der Grundgleichung. Hilfsbegriffe. Arbeit, Energie, Impuls. Drehbewegung fester Körper. Beschleunigte Bezugssysteme. Einige Eigenschaften fester Körper. Über ruhende Flüssigkeiten und Gase. Bewegungen in Flüssigkeiten und Gasen · Akustik: Schwingungslehre. Fortschreitende Wellen und Strahlung · Wärmelehre: Grundbegriffe. I. Hauptsatz und Zustandsgleichungen idealer Gase. Reale Gase und Dämpfe. Wärme als ungeordnete Bewegung. Transportvorgänge, insbesondere Diffusion. Die Zustandsgröße Entropie. Umwandlung von Wärme in Arbeit. II. Hauptsatz · Tafeln · Sachverzeichnis.

AUS BESPRECHUNGEN FRÜHERER AUFLAGEN

„Jeder Physiker sollte sich einmal mit dem Pohlschen Buch auseinandersetzen, das eine Fülle von Anregungen bringt ... Für Anfänger ist das Buch allerdings nicht unbedingt geeignet, dafür um so mehr für jeden, der tiefer in die behandelten Gebiete eindringen will." *Der Mathematisch-naturwissenschaftliche Unterricht*

SPRINGER-VERLAG · BERLIN · GÖTTINGEN · HEIDELBERG

An introduction to the geometry of numbers

By J. W. S. Cassels
Fellow of Trinity College, Cambridge/Great Britain

(Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete. Herausgegeben von R. Grammel, F. Hirzebruch, E. Hopf, H. Hopf, W. Maak, W. Magnus, F. K. Schmidt, K. Stein, B. L. van der Waerden. Band 99.)

In englischer Sprache. With 12 Figures. VIII, 344 Seiten Gr.-8°. 1959.
DM 64,50; Ganzleinen DM 69,—

CONTENTS

Notation · Prologue · Lattices: Introduction. Bases and sublattices. Lattices under linear transformation. Forms and lattices. The polar lattice · Reduction: Introduction. The basic process. Definite quadratic forms. Indefinite quadratic forms. Binary cubic forms. Other forms · Theorems of Blichfeldt and Minkowski: Introduction. Blichfeldt's and Minkowski's theorems. Generalisations to non-negative functions. Characterisation of lattices. Lattice constants. A method of Mordell. Representation of integers by quadratic forms · Distance functions: Introduction. General distance-functions. Convex sets. Distance functions and lattices · Mahler's compactness theorem: Introduction. Linear transformations. Convergence of lattices. Compactness for lattices. Critical lattices. Bounded star-bodies. Reducibility. Convex bodies. Spheres. Applications to diophantine approximation · The theorem of Minkowski-Hlawka: Introduction. Sublattices of prime index. The Minkowski-Hlawka theorem. Schmidt's theorems. A conjecture of Rogers. Unbounded star-bodies · The quotient space: Introduction. General properties. The sum theorem · Successive minima: Introduction. Spheres. General distance-functions. Convex sets. Polar convex bodies · Packings: Introduction. Sets with $V(S) = 2^n \Delta(S)$ Voronoi's results. Preparatory lemmas. Fejes Tóth's theorem. Cylinders. Packing of spheres. The product on n linear forms · Automorphs: Introduction. Special forms. A method of Mordell. Existence of automorphs. Isolation theorems. Applications of isolation. An infinity of solutions. Local methods · Inhomogeneous problems: Introduction. Convex sets. Transference theorems for convex sets. The product of n linear forms · Appendix · References · Index.

SPRINGER-VERLAG · BERLIN · GÖTTINGEN · HEIDELBERG

Springer-Verlag, Berlin · Göttingen · Heidelberg. — Druck der Universitätsdruckerei H. Stürtz AG., Würzburg
Printed in Germany